Assessing Mathematics Proficiency of Multilingual Students: The Case for Translanguaging in the Democratic Republic of the Congo

FERNANDA GANDARA AND JENNIFER RANDALL

Monolingual content assessments of multilingual students remarkably fail their most essential endeavor—to provide meaningful information about their content proficiency. Multilingual students take years to catch up with their monolingual peers in terms of language proficiency, and it is inappropriate to assess them using a language that they do not fully understand. Because multilingual students do not behave as multiple monolinguals, translated tests are not a satisfactory solution. Test developers should produce assessments that enable multilingual students to use their entire linguistic repertoires and engage in their natural linguistic practices. This study examines a translingual administration of a mathematics assessment in the Democratic Republic of the Congo. The findings show that a translingual administration is more appropriate than a traditional administration. The findings, however, raise doubts about standardization and call for further deliberation on how to properly assess mathematics in similar contexts.

Educational institutions constantly make language choices that affect schools, as well as the communities and students that they serve. For example, schools must choose which language to use for instruction of some or all of the subjects, for some or all of the grade levels, a decision that necessarily involves the larger and complex sociolinguistic reality in which the school is situated. From a pedagogical perspective, language decisions need to be selected to guarantee the optimal preparation of students for life and citizenship, yet language policies in education do not always match these considerations. Research shows that strong forms of bilingual education—that is, programs where the aim is that students become bilingual, regardless of whether or not they serve language minorities or majorities—are preferred in terms of learning outcomes (García 2009; Baker 2011) and are more adequate to serve multilingual groups of students who typically take at least 5 years to catch up with their monolingual peers in terms of language proficiency (Shohamy 2006, 2011; Baker 2011). In reality, however, we note a scarcity of strong bilingual education programs in many parts of the world, including Sub-Saharan Africa, partly due to ideological reasons, because language is a pow-

Received December 20, 2017; revised October 1, 2018; accepted October 3, 2018; electronically published January 3, 2019

Comparative Education Review, vol. 63, no. 1.

0010-4086/2019/6301-0044$10.00

erful tool to strengthen national identity (Shohamy 2006; Tollefson 2008). Overall, the contrast between research and practice with regards to language decisions suggests that in education, politics often proves stronger than research (García et al. 2006; Shohamy 2006; Baker 2011) and that language choices are far from solely a pedagogical matter.

Language decisions that are not pedagogically oriented are problematic because they constrain and may even deprive groups of students from the opportunity to learn. The language practices of schools contain messages and establish hierarchies that affect students' identities (García et al. 2006; Shohamy 2006; Tollefson 2008; García 2009) and may bring detrimental consequences for self-esteem and well-being if some students feel neglected, stigmatized, or left out. Monolingual lenses in education—which value and promote only one language, often that of the mainstream (Garcia 2009; Baker 2011)—exclude or delegitimize linguistic diversity, leading to the marginalization of language minorities and the creation or maintenance of social inequalities (García et al. 2006; Shohamy 2006). In a world where multilingualism is the norm and one that is characterized by massive social movements and displacement of people, any attempt to improve equitable access to education needs to critically examine linguistic policies of schools and educational systems.

Often overlooked is the language of educational assessments or tests.[1] Tests affect students at many stages in their educational journey and are used for several purposes, such as diagnosis of strengths and weaknesses, placement decisions, or for accountability (Sireci and Gándara 2016). Test developers make numerous language decisions, including the language of the questions, the language of instructions, the language in which students may respond to the assessment, and the way in which to penalize language mistakes in scoring responses. These language choices shape behaviors and set expectations for students, expectations that cannot be overlooked, since testing may further enforce educational language policies (Shohamy 2006, 2011). The language of tests should be carefully evaluated.

One of the strongest criticisms toward educational tests emerges from the fact that many of them—at least standardized assessments—do not recognize the wide linguistic variety of test takers. Standardized assessments are typically *monolingual* in their focus (Escamilla 2006; Shohamy 2006; García 2009; López et al. 2017). In other words, these assessments are developed as if the language of test takers is homogeneous and standard, neglecting the reality that *emerging bilinguals* (students who through schooling and by acquiring the language of the country they live in become bilingual; García et al. 2008) and more generally *multilingual students* (students who are exposed to

[1] Assessment and test are used interchangeably.

multiple languages in their developmental process and have some level of proficiency in more than two languages) encounter challenges in their classrooms, such as using language flexibly (Creese and Blackledge 2010). Monolingual tests are also questionable on ethical grounds as they explicitly favor ideologies that seek integration, cohesiveness, and sociopolitical problems by means of enforcing the language of the majority. Therefore, they are hardly appropriate to address the challenges and complexities that are inherent to multilingual educational contexts.

Problematizing Language

The power of languages as a form of control and hegemony has been largely documented, but acknowledging the connections between language and power is insufficient to overcome the problems that it raises. In educational debates, language can be treated as an autonomous system, as something that exists outside and above human beings. Makoni and Pennycook (2007) bring light to the flawed nature of this conceptualization and argue that not only are languages not autonomous systems but that they are actually inventions. Languages as autonomous systems are incongruent with sociolinguistic approaches that focus on language ideologies: if comprehending language use requires acknowledging the beliefs of the participants, we simply cannot state that language is independent of its context. Cogently and extensively, Makoni and Pennycook (2007) suggest that a necessary step in uncovering and undoing the political power of languages is to come to good terms with the idea that language is an invention. To understand the metadiscursive, ideological, and historical origins of such inventions is urgently needed to overcome the current limitations in linguistic debates.

The idea that language is a social construction is well supported, for example, by the lack of consensus on the number of languages (Makoni and Mashiri 2007; García 2009). The social invention of languages is particularly evident when studying colonialism, a time in which language was used to carry out the agendas of missionaries and colonial administrators. To fulfill their purposes, both missionaries and colonial linguistics embarked on a project of creating languages: constructing grammars, orthographies, dictionaries, even language academies (García 2009). Language as an invention does not mean that people did not have language before colonialism, but that the notion of language and the characteristics of the "languages" that were used to describe, define, and control colonized populations were created after European, Western, Colonial, and Christian ideologies and metadiscursive regimes. Missionaries and colonialists created languages that resembled their own views and reproduced their own semiotic systems: from the words that were added to the dictionaries (e.g., words related to the divine) to the efforts to construct familiar grammatical structures, missionaries

failed to acknowledge that languages are a subjective human construct and created languages that were no one else's mother tongue (Makoni and Mashiri 2007).

Despite their invented nature, the effects of language are very real (Makoni and Pennycook 2007). Conventional ideas about languages or about language continuously oppress certain groups of the population. For example, Branson and Miller (2007) explain how traditional approaches to analyzing a language known as *Kata Kolok* (sign language) is detrimental to their speakers, as *Kata Kolok* does not meet the Western and academic criteria of what a language should look like. In particular, *Kata Kolok* does not meet the criteria of autonomy postulated by traditional linguistics, as its signs are not arbitrary (Branson and Miller 2007). Therefore, any traditional analysis of *Kata Kolok* is likely to result in some form of epistemic violence or oppression. As a response, researchers like Makoni and Pennycook (2007) call for a disinvention of our notions about language. Understanding that languages are socially and politically constructed is essential to face situations in which there are reasons to change them or the way we think about them. Concretely, the ideology of invention creates an additional layer of complexity to any analysis of language in education.

Problematizing the Language of Mathematics Assessments

Mathematics content is delivered through language (Adoniu and Quing 2014); when testing one's knowledge of mathematics, one is partly testing the language of mathematics. This connection creates important challenges for test developers as the linguistic demands of mathematics are high and should not be overlooked. For example, mathematics demands learners to master symbolic, graphic, tabular, formal, and technical language, and to be comfortable with multiple representations (National Governors Association 2012). Moreover, the structure of math sentences may be different than that of the language of instruction and may rely heavily on particular grammatical structures such as dependent and relative clauses (Adoniu and Quing 2014). Mathematics also requires students to acknowledge that certain words have different meanings in math than outside the subject (e.g., table meaning "times table," or "integration" referring to a calculus operation). The polysemy found in mathematics can cause difficulties for students, particularly for linguistic minorities (Martinello 2008; Adoniou and Qing 2014) and multilingual students. And all of these challenges tend to increase as students progress through grade levels.

To properly identify the linguistic demands of mathematics, some researchers propose to use Cummins' (2000) concept of *academic language* as a framework to identify and operationalize these linguistic demands. Academic language refers to language that pertains to academic settings, which is

less contextualized and more abstract than *social language*, contextualized "everyday" use of language (Cummins 2000). Academic language can be described from a vocabulary, a grammatical, or a discourse function (DiCerbo et al. 2014) and may clarify which characteristics of mathematics language are relevant to the subject and which characteristics are redundant. Using this framework to develop mathematics tests has gained support because academic language is a significant predictor of students' proficiency across several content areas (Bailey 2007). However, it is not clear whether practitioners will easily be able to disentangle academic language from social language, especially because there is no single view around the language of mathematics (Brown 2002). More importantly, the concept of academic language falls short in addressing the sociopolitical challenges that testing mathematics entails. The academic language of mathematics can be presented as neutral or objective, yet this is exactly the type of discourse that needs to be challenged.

Multiple tensions are present in mathematics teaching and learning in multilingual contexts. These tensions include the dilemma of code-switching, the dilemma of mediation, and the dilemma of transparency (Adler 2001; Barwell and Pimm 2016). In multilingual classrooms, mathematics teachers are faced with specific linguistic challenges such as which language(s) to use for teaching, how to mediate understanding when learners bring multiple languages and levels of proficiency, how to give students voice without straying from the conversation, where to set the boundary as to what is a proper form of mathematical discourse in terms of meaning and form, and so forth. Teachers sometimes prefer one language over others for political reasons. In interviews, Setati (2008) found that all teachers (six in total) indicated that they prefer to teach mathematics in English because of its status as international language (ideological reasons) or the possibilities that it opens up for students (pragmatic reasons). Barwell et al. (2016) also note that most of the teachers interviewed in Canada, South Africa, and Malaysia prefer to teach in mainstream languages such as English or French despite the multiple tensions this decision entails. In contrast, both Setati (2008) and Barwell et al. (2016) describe that certain stakeholders—either learners, teachers, or communities—prefer mathematics to be taught in a local language to ensure understanding. These stakeholders correctly believe that students learn best when taught in a language that they understand, a finding that has been echoed by multiple organizations and researchers (e.g., Global Education Monitoring Report 2016). However, in multilingual contexts there are many confronting cultural models (Gee 1999) in place, and language choices are entrenched within a multidimentional space that is far from solely pedagogical. To be sure, the language of mathematics is complex, always in tension, and the challenge of making adequate lingusitic choices when developing mathematics tests remains high.

The Promise of Translanguaging

The traditional solution to deal with the inappropriateness of monolingual tests has been to use *linguistic accommodations*. Linguistic accommodations are changes to the language of the test via adjusting the materials or procedures, with the aim to increase accessibility to the content of the test or reduce barriers for linguistic minorities (Thurlow and Kopriva 2015). Linguistic accommodations include the provision of dictionaries, the translation of instructions or item stems, and reading the questions aloud, to name a few. Yet linguistic accommodations may not be entirely appropriate since they typically operate on the assumption that there is redundant linguistic complexity that can be simplified and that they are needed on a temporary basis (Abedi 2004, 2009). A more adequate approach would be to develop tests that acknowledge the linguistic diversity of their test takers and that do not have eventual expectations of monolingualism on the side of the examinees. Research shows that it takes a long time for second language learners to achieve the language proficiency of monolingual peers (Cummins 2000; Abedi and Gándara 2006; Shohamy 2006), even longer for students who come from disadvantaged backgrounds (UNESCO 2016). The time is even longer in the context of mathematics (Shohamy 2006, 2011) because math proficiency requires mastering both content and a second language. Likewise, there is evidence that students continue to use their first languages for a long time in academic situations (Shohamy 2006, 2011). Temporary and partial solutions to the language problems of educational tests are unsatisfactory, as multilinguals do not behave and will not behave as one, two, or multiple monolinguals (García 2009; Shohamy 2011).

A different approach to address the linguistic shortcoming of assessments of multilinguals is rooted in the concept of *translanguaging*, as both a practice and a lens. Translanguaging refers to the language education pedagogy in which students' receptive language is different to their productive language (García 2009). More generally, translanguaging as a practice refers to the flexible use of linguistic resources that characterizes bilinguals in their attempt to make sense of their bilingual worlds (García 2009; Creese and Blackledge 2010). Translanguaging as a lens does not conceive multiple languages as separate but rather conceptualizes the language of multilinguals as an integrated system (Canagarajah 2011; Velasco and García 2014) and assumes that what makes human communication unique is selecting features from the entire linguistic repertoire to produce an intended message. The focus of translanguaging—the lens—is not on languages but in the way in which bilinguals use their linguistic resources. Importantly, as some have pointed out, translanguaging is not the same as code-switching. Code-switching arises from an external view on languages, one that considers that multilinguals switch between languages that exist as autonomous systems

(García et al. 2017). In contrast, translanguaging focuses on the internal perspective of a student who flexibly selects features among all those that he or she possesses.

Translanguaging provides an alternative framework to examine and solve the problem of assessments for emerging bilinguals, particularly in mathematics. Translanguaging is supported on an empirical basis, as it reflects common behaviors encountered in the bilingual classroom (Creese and Blackledge 2010; Canagarajah 2011). A translingual framework could therefore enhance the alignment between assessment and instruction, something that is rarely thought of from the linguistic perspective. Finally, translanguaging departs from the notion of languages as structured, fixed, and standardized systems of signs, overcoming the epistemological problem of treating language as an autonomous system. In brief, translanguaging is a promising approach to reimagine the assessments of multilinguals.

Using translanguaging in mathematics assessments requires normalizing multilingual competence (García 2009) by incorporating linguistic flexibility at the level of questions and responses, in terms of language and modalities. This is conceptually different than using direct linguistic accommodations, as translanguaging neither assumes redundant linguistic complexity nor works under expectations of eventual (double or multiple) monolingualism. To properly accommodate the practices of every student, this linguistic flexibility has to be self-regulated and dynamic. Students need to be able to activate their entire set of resources and deploy them according to their individual preferences. We do not need to provide students with "access" to the content of the test but rather create tests that are built considering the language practices that bilinguals use to make meaning and to communicate. Finally, a translingual approach to mathematics assessments also requires incorporating opportunities for students to interact with others (López et al. 2017); translanguaging takes place in social interactions, and tests should include some degree of interactivity.

Purpose

When assessing the knowledge and skills of emerging bilinguals or multilinguals, using monolingual tests is a flawed and inappropriate practice at best. Monolingual tests of mathematics make it difficult to determine the extent to which the performance of bilinguals is a matter of content knowledge and/or a matter of language proficiency (Abedi 2004; García 2009; Shohamy 2011). Translanguaging offers a new framework to think about mathematics assessments in multilingual contexts, a framework that has received increased attention and support from researchers, and one that may better serve the needs of multilingual communities than approaches rooted in the concept of academic language. However promising, there is no con-

sensus as to whether translingual approaches to assess mathematics are appropriate. To that end, the purpose of this study was to evaluate the appropriateness of the translingual administration of the Early Grade Mathematics Assessment (EGMA) on a group of bilingual girls in the Democratic Republic of the Congo. By translingual administration we mean one that incorporated linguistic flexibility at the level of instructions, questions, and answers, and that incorporated an element of interactivity. The emphasis on the word administration is to state that the test was not developed using a translingual framework: it is only the administration that was reconceptualized accordingly. The research question that guided this work is as follows: is the translingual administration of the EGMA appropriate for the context in which these data were collected, as evidenced by (*a*) the characteristics of the context in which the assessments were administered, (*b*) the level and ways in which girls engaged in flexible bilingualism, and (*c*) the alignment with teachers' practices and beliefs around how to assess math proficiency?

Method

The Sample

This qualitative study was conducted within a larger 4-year randomized control trial (RCT) evaluation study designed to measure the impact of a package of educational interventions on the learning and attendance outcomes of primary school girls. The intervention—an initiative through the Girls' Education Challenge—targeted 100,768 marginalized girls in the Democratic Republic of the Congo (DRC).[2] Participants were selected from four schools within the study in the province of Equateur (city of Mbandaka). In each of the four schools, between 30 and 40 girls from grades 3 through 6 were randomly selected. After selecting the girls, approximately half of them were randomly assigned to the translingual EGMA, and the other half were assigned to the traditional EGMA administration. With the traditional administration, girls had the test administered by one enumerator and could choose only one language among two for instructions and responses. Here we focus on the translingual administration and our sample consisted on the remaining girls who received the translingual treatment. Table 1 displays the total number of students included from each grade. Teachers who participated in the study taught math in grade levels 3 through 6, in one of the four schools sampled.

The Context

The EGMAs took place in the urban area of Mbandaka, but the label of *urban* does not translate into the concept imagined from high-income

[2] Because the larger study focused on improving the learning outcomes of girls, the sample is composed solely of girls. Boys were not assessed.

TABLE 1
GRADE LEVEL DISTRIBUTION OF GIRLS

Translingual EGMA	3rd	4th	5th	6th	Total
Total number of students	22	13	33	12	80

countries. Mbandaka has some paved streets, but most of them are not, and its downtown is composed of a few streets with small businesses: while remarkably crowded and dense and therefore technically a city, Mbandaka shares many characteristics of rural regions in other parts of the world. Some of the schools that we visited were located miles away from downtown, in rather isolated communities. The weather conditions were extremely hot, dusty, with unexpected and dramatic rains. It is important to grasp that schooling conditions in this area are very difficult, because students must deal with harsh weather on top of the challenges and injustices that poverty entails.

The Assessment

Traditional Early Grade Mathematics Assessment: The EGMA is a test that was designed to reliably measure early grade mathematical skills across a wide set of countries (Platas et al. 2014) and is widely used in the context of international education. It is orally and individually administered and has to be locally adapted. This adaptation typically involves selecting the final subtasks and developing the test in local languages. Ultimately, students are allowed to choose one language of administration. The EGMA (Reubens 2009) used included five subtests: reading/recognizing numbers ($\alpha = .940$), comparing quantities ($\alpha = .861$), number sequences missing values ($\alpha = 0.772$), addition ($\alpha = .911$), and subtraction ($\alpha = .925$). The traditional administration of the EGMA typically takes about 15 minutes per child to administer, all of the subtests are orally administered, and some subtests are timed.

Translingual Early Grades Mathematics Assessment: The translingual administration of the EGMA followed the two-step approach proposed in López et al. (2017), which allows students (*a*) to draw from their entire repertoires (i.e., use all their languages) and (*b*) to engage in the interactions that stimulate the creative and meaningful use of linguistic resources. Two trained enumerators, including the first author, administered the test and delivered the instructions in French, the official language of instruction, and/or Lingala, the lingua franca of the province. Some instructions were delivered in French, some in Lingala, and some were delivered in both. There was no prescribed pattern to use language in instructions, and each administration was unique in this regard. The only standard condition was that all girls were exposed to a mix of languages and that the enumerators

made optimal choices based on the three-way interactions between the enumerators and the student. To support the enumerators, sample instructions were printed in both French and Lingala. Second, the girls were asked to give their responses in French, Lingala, or both.

The reader should note that because this study was a part of a larger evaluation study (specifically an equating study), girls were administered two forms of the EGMA. The primary purpose of this larger equating study was to ensure that previous and current forms of the assessment were indeed equivalent for the purposes of making inferences about girls' math literacy over the 4-year period.

Procedures and Analysis

We relied on three sources of data for analysis and interpretation: (1) field notes taken during test administrations, (2) tape recordings of test administrations, and (3) semistructured interviews with teachers. Each source of qualitative evidence was collected to capture the characteristics of the test administration, the ways in which the examinees used language during the test administration, and/or the ways in which (or the extent to which) the translingual administration of the test aligned with teachers' own teaching and assessment practices. Researchers took field notes for all 80 administrations of EGMA and randomly tape recorded half of the administrations. As part of the analysis, we conducted an interpretation analysis of the field notes with the aim to recover the essential characteristics of the administrations and to evaluate the appropriateness of the translingual EGMA in relation to the context. Additionally, we conducted a structural analysis of the students' test speech using the 40 tape recordings. This analysis was meant to uncover the ways in which girls used language during the EGMA administrations and to evaluate whether the translingual EGMA was appropriate given their linguistic patterns. Two themes were explored: (*a*) the extent to which girls engaged in flexible bilingualism and (*b*) the characteristics of girls' flexible bilingualism. The nature of this analysis was descriptive.

Finally, we conducted semistructured interviews with the eight teachers at these four schools. The interviews were conducted in French, tape recorded, and took no longer than 5 minutes each. The objectives of the interviews were (*a*) to broadly understand teachers' views on how to evaluate their students' math knowledge and (*b*) to gain insight on the degree of alignment between the translingual EGMA and their views and practices in the mathematics classroom. The aim was to gain a better understanding of the evaluation practices that teachers engaged in the classroom and evaluate the extent to which the translingual EGMA was appropriate. The analysis of the interviews was interpretive and consisted of identifying a set of themes that could inform both objectives described in the above text.

Results

To respond to our research question, we focused on three different themes: appropriateness regarding the context, appropriateness regarding the way in which students used language, and appropriateness regarding teacher beliefs.

Appropriateness regarding the Context

The most salient theme across the field notes relates to the inherent difficulty that this context posed to any standardized assessment: enumerators could not effectively deliver the assessment in a standardized way. The translingual administration did not aim for standardization at the level of language, but it did aim for some standardization at the level of conditions (e.g., silent room, no interruptions, same amount of time per task, etc.). As we administered more and more assessments, it was difficult to envision a standardized administration for the assessment even at the level of essential conditions. The use of standardized assessments, in the way that it is typically conceptualized/instituted in western contexts, warrants reexamination within the context of Mbandaka. Some of the notes in this regard are as follows:[3]

1. "None of the instructions were repeated exactly in the same way, and it feels more natural and appropriate for this context."
2. "Interruptions are so normal, they [other enumerators] do not even realize the effect that this has on standardized assessments."
3. "This administration is as standardized as possible."
4. "Imposible estandarizar instrucciones." [English: It is impossible to standardize instructions.]
5. "Standardization is def. [definitely] not useful in this context."

One of the core characteristics of the translingual administration was the variety with which the EGMA instructions were delivered across administrations, yet this lack of standardization extended to the highly irregular ways in which enumerators used the chronometers for timed responses. Some of the notes that highlight this issue are as follows:

6. "The enumerator easily forgets that she has to time the exercise."
7. "It is worth to be flexible w/th 5 secon. [with the 5 second rule]."
8. "Sonó un celular: estos retrasos de tiempo {o por ejemplo que las niñas afuera estén molestando} no se incluyen en el tiempo de administración. Confiar en la estandarización de este instrumento es difícil. Me parece que la administración oral y local lo hace imposible." [English: A

[3] All notes correspond to the first author and are written in first person.

cell phone rang: these delays (or, for example, that girls are outside distracting) are not included in the time of administration. To trust the standardization of this instrument is difficult. It seems to me that the oral and local administration makes this impossible.]

9. "I do not stop them if they are en train de donner les responses." [English: I do not stop them if they are in the process of giving the answers.]

The irregularity that characterized the instructions and the timings, which were delivered by trained enumerators who had administered several EGMAs or tests like the EGMA in their lives, seemed to index a larger cultural phenomenon: irregularity being an essential characteristic of Mbandaka. We neither counted the number of interruptions nor classified the different types of disruptions, but we had many pauses due to unexpected events. We also had to change plans for many occasions, for similar reasons. This lack of standardization was reflected more subtly in other behaviors. For example, it was difficult for enumerators to remember to turn off their cellphones while administering the assessment (an explicit requirement of the EGMA administration), which often rang. This relative unconcern toward the regularity of the administration showed that enumerators did not internalize the concept of standardization or its relevance for the assessment. Regarding the girls, while all of them were administered the assessment twice, it was rare for them to recall the instructions during their second EGMA:

11. "Girls are as surprised with instructions as the first time."
12. "We repeat the instructions every time, and it sounds strange to me to say the exact same thing."
13. "I tried to go without saying the instructions, but she didn't remember. From 10 minutes ago!"
14. "She didn't understand, so we had to repeat instructions and add explanation. This happens a lot."

In sum, the impossibility of standardization in this context suggests that flexible assessments are more appropriate in contexts like Mbandaka. In that regard, the translingual EGMA is more appropriate than the traditional EGMA, where instructions are given in one language only and in a fairly standardized manner. However, the analysis of the field notes and of the administrations themselves suggests that standardization has to be negotiated at multiple levels, not just around language.

Appropriateness regarding Language Use

The translingual EGMA was not ideal to explore girls' speech, since answers were too short to gain a thorough notion of their test speech. There

were only two tasks in which girls could use (and typically used) sentences to provide their answers: number comparison and missing numbers. In the administration of these tasks, we encouraged girls to respond using sentences rather than single numbers (e.g., "The missing number is X"). However, even when girls used sentences to provide answers, these sentences were incredibly bounded. Consequently, we did not conduct a speech analysis, but we identified general trends within the use of language when communicating math knowledge.

Based on the test speech of 37 girls,[4] we classified their responses according to their lexical composition: French-only responses, Lingala-only responses, and mixed responses. This analysis was done at (*a*) a test level (both forms), (*b*) the form level, and (*c*) the task level. Regarding the test level, 70% of girls used words from more than one language to respond to both forms of the translingual EGMA. Interestingly, girls who used a single language to provide their responses used French-only responses.

The second analysis of responses examined the test speech across the two forms. Most girls showed stable patterns of lexical use across forms, yet 29% of the tape-recorded girls changed the way in which they used language on the second translingual EGMA administration. Importantly, there is directionality in the switch: 9 of the 10 girls who used language differently across forms, deployed mixed language in the first and French-only in the second, which suggests that translanguaging may have acted as a scaffolding device.

More insightful results appear when we examine this variation at the task level. Table 2 shows that most of the differences took place with the "wordy" tasks—number comparison and missing number. In other words, when girls changed the way in which they delivered their responses (e.g., from Lingala only to French only), they did so with higher frequency for those tasks that entailed more elaborated responses.

Table 3 shows the number of girls who responded to each task using Lingala only, French only, or Mixed languages responses. We note from table 3 that there are three tasks where students rarely or never gave responses using words from more than one language: number identification, addition, and subtraction. We also note that Lingala-only responses took place twice only, suggesting an implicit internalized hierarchy between French and Lingala.

When we analyze the responses for the number comparison and missing number tasks, we observe a stable pattern of language use: numbers are mostly expressed in French, and operations or comparative language are mostly expressed in Lingala. Almost every girl who used mixed language to respond to these items followed that pattern (see, e.g., excerpt 1). It is unclear to what degree the translingual administration influenced this out-

[4] Three audio recordings were unintelligible.

TABLE 2
LEXICAL ITEMS ACROSS FORMS BY TASKS

	Same	Different
Number identification	34	1
Number comparison	26	9
Missing number	27	8
Addition	31	4
Subtraction	33	2

come. To be clear, there was no statistical relationship between the language of the instructions and the language of the responses, yet some tape recordings suggested a relationship between the way in which instructions were explained and the responses provided by the girls.

Excerpt 1: Examples of Responses to Each Task (Lingala in Bold)

Number comparison.

- **Oyo eleki** trente-huit (this [number] exceeds 38).
- **Oyo eleki** cent cinquante-quatre (this [number] exceeds 154).

Missing numbers.

- **Awa** neuf **ezangi** (here nine [is] missing).
- **Moko** douze (the number is 12).

Based on these results, we conclude that the translingual EGMA is certainly more appropriate than a traditional EGMA where only one language is used. First, even in a short and low-speech assessment like the EGMA, most children deploy multilingualism when allowed. It is simply inappropriate to ask this population to express themselves in one single language. Second, we do observe some stable linguistic patterns. For example, the fact that students who switch from a mixed response (form 1) to French-only responses (form 2) may do so because French is the official language of instruction from the third grade, and they assume that French-only responses is the expectation. Translanguaging assessments may undo this monolingual expectation, a result that we consider appropriate. Also, the fact that some of the observed linguistic patterns (including the directionality in lan-

TABLE 3
LEXICAL ANALYSIS BY TASKS ACROSS BOTH ADMINISTRATIONS

	Number Identification	Number Comparison	Missing Number	Additions	Subtractions
Lingala	1.00	.00	.00	1.00	.00
French	70.00	39.00	46.00	67.00	69.00
Mixed	.00	33.00	25.00	3.00	2.00

The numbers do not add to 74 as not all students answered all tests/items.

guage switches) were relatively stable and are aligned with the way in which bilingualism is enacted in the DRC, deserves more deliberation. For instance, the way in which students provided their responses, French numbers, and Lingala explanations seem to follow a reality where teachers attempt to teach mathematics in French yet use Lingala to enhance understanding, a pure form of translanguaging. This pattern is also consistent with the views of mathematics teachers and students reported by Setati (2008) and Barwell et al. (2016), where political and pedagogical objectives are in confrontation. Translingual assessments of mathematics would be better aligned to the way in which math is truly delivered, at least in the context of Mbandaka, and partly reconcile the divide between politics and pedagogy. Finally, the fact that girls used a higher proportion of mixed responses for the "more wordy" tasks suggests that longer responses creates more opportunities for girls to use language flexibly. Translingual assessments of mathematics would therefore be especially appropriate for tasks that involve more speech, such as problem solving.

Appropriateness regarding Teachers' Beliefs

To analyze teachers' beliefs regarding the assessment of mathematics, we focused on three themes: preferred method of evaluating mathematics proficiency, use of the evaluations, and views around the language in which to teach mathematics. Regarding their preferred method of evaluation, almost all teachers selected traditional assessments such as oral interrogation of students, homework, exercises, and/or tests, which were mostly taken from or developed after textbooks. Each teacher mentioned at least two methods, suggesting some variety in the evaluation methods they use. In addition, most teachers believed that the responses to these assessments were the best way to detect whether a student had learned "a knowledge" or not. There was almost perfect consistency between their preferred method of evaluation and the way in which they came to realize whether students were learning or not.

Regarding the use of the evaluations, there is no clear picture of the common practice: some teachers expressed using them to make instructional decisions, some others complained about not being able to make those decisions, and a few teachers mentioned that they used evaluations to assign students to remediation lessons. Our interpretation is that teachers are willing to use assessments/evaluations formatively, yet this is far from a consistent practice. Most teachers mentioned how they changed their "strategy" based on assessment results, mostly meaning that they changed the way in which they taught the content. However, it was clear from the analysis of the different interviews that there was some variation in the way in which teachers used the term strategy. For instance, one teacher used the

word "strategy" to mean "focus": to him, changing strategy meant to address the content that students should reinforce. There is not enough information to elaborate on this semantic difference, yet it is important to mention that even at this level of discourse, there was semantic variation between the teachers, and this does speak to the way in which people from Mbandaka use language.

Regarding the language of instruction, none of the teachers said that they used one single language in the classroom. Some teachers began saying that French was the language of instruction, but that students did not fully understand French or could not fully express themselves in French; therefore, whenever needed, they used Lingala to explain to ensure that students understood the content. Other teachers began by saying that they used both languages for instruction but, after some iterations, changed their answer and said that while they used both French and Lingala, they typically used Lingala to ensure better comprehension. Other teachers said that they used both languages, French and Lingala, in "pure" or mixed versions and that students were comfortable with both languages.

Teachers' responses suggest that all students understand Lingala, and that this is the language in which they learn better. French is the preferred language of instruction, the "recommended" language—as one school director said in a side conversation—yet it cannot be used in isolation, as students would not understand the content. It is also interesting to note that all of the teachers referred to French by its name, but not all of them referred to Lingala by its name. Lingala was referred to as "the mother tongue," the "local language," and "our national language." This reflects how speakers of Lingala hold different beliefs toward their language.

In brief, teachers' responses provide support for a translingual version of the EGMA. On one hand, teachers value traditional methods of evaluation, and tests such as the EGMA meet that criteria. Second, teachers use language flexibly in the mathematics classroom, meaning that a translingual EGMA improves the alignment between assessment and instruction. Third, the lack of standardization in the way in which teachers used language and/or referred to language suggests that flexible assessments are more appropriate because they provide more opportunities to clarify what is the task at hand. Finally, teachers seem willing to use assessments formatively, and translingual assessments may improve the quality of formative data as they enhance opportunities for students who do not necessarily master the language of instruction (see, e.g., Ascenzi-Moreno 2018.)

Discussion

This article examined the appropriateness of a translingual administration of a standardized mathematics assessments, via several analyses of

field notes, tape recordings, and interviews. We first noted that the context in which the administrations took place—the city of Mbandaka in the Democratic Republic of the Congo—presents high barriers to standardized assessments as standardization itself seems to be an exogenous concept. The value of standardized assessments—assessments that are administered and scored in the same exact manner for the entire group of test takers—is rooted in the comparability of results, one of the main purposes of achievement tests (Popham 2016). To some, especially in the West, standardization is natural in educational contexts, but this assumption is challenged in contexts like Mbandaka. Irregularity was observed in every step of the translingual EGMA administration raising doubts about its appropriateness. Indeed, even in the West the appropriateness of standardized assessments has repeatedly been called into question (see Rotberg 1995; Popham 2001, 2003, 2008) as their use has been linked to (*a*) a culture of curriculum reductionism, (*b*) adverse impacts on students' self-perceptions, as well as (*c*) student- and teacher-level cheating. Wells (2014) has argued that the current so-called color-blind accountability practices in the United States, which rely primarily on the use of standardized testing as the measure of achievement are, in fact, not color-blind at all and provide a disservice to all students. Perhaps most notably, critics have noted that standardized assessment systems can be misinterpreted as meritocratic assessment systems both perpetuating the illusion of objectivity and masking sociocultural disparities in the United States and abroad (see Goodman 2013). Still, given that low levels of enacted standardization may lead to misleading conclusions and that standardization is central to the validity argument of tests like the EGMA, further research should question the extent to which evaluators can assume standardization of other administration conditions, such as equally timed tasks or no-interruptions, in regions like Mbandaka. A priori and based on this experience, in contexts like Mbandaka it seems desirable to move away from such assumptions and to reformulate the validity argument of the tests used to evaluate large-scale projects.

Second, a large majority of girls used words from more than one language to provide their responses. In addition, almost one third of the girls responded to a same task differently across forms. This occurred mostly for the "wordy" tasks—number comparison and missing numbers, which were also the tasks where girls deployed higher levels of bilingualism. These results are outstanding, as the EGMA test is a low-speech test, to the extent that all tasks could be answered with single numbers. Yet most of the girls used language flexibly at the task, form, and test level. This evidence strongly indicates that in Mbandaka, any monolingual or double monolingual (e.g., pick a language among these two) approach to testing is inappropriate, even when the language of the test is a so-called mother tongue language. Bilingualism was the norm among these girls; therefore, the

translingual EGMA is better suited to serve this context than a traditional administration of the test.

Third, teachers declared using various evaluation methods, which are fairly traditional. The majority of interviewed teachers said that they used exercises from the instructional material that they possessed to create homework or tests. These findings suggest that—in this context—classroom evaluation practices are largely influenced by the instructional materials that are provided. This is consistent with international evidence that textbooks convey more than knowledge and can foster change beyond content proficiency (e.g., UNESCO 2017). If translanguaging is appropriate to evaluate the content knowledge of multilingual students, a way in which to foster translanguaging in assessments is by developing and providing heteroglossic instructional materials to the teachers. There is evidence that heteroglossic materials, with diversity of discourses, languages, and voices, can meet pedagogic expectations while counteracting separatist and hegemonic linguistic practices (Busch and Schick 2007). Instructional material should include heteroglossic methods to assess content knowledge, a move that would certainly open up the possibilities of using translanguaging in assessment in contexts like Mbandaka.

Overall and taking all these findings into account, we can state that the translingual EGMA is more appropriate than a traditional administration. Translanguaging seems to be a common practice in the Mbandaka classroom, used to negotiate the understanding of students who cannot fully engage in a French-only lesson. From the way in which teachers talked about Lingala, it is likely that translanguaging is also used as a means to perform identities in the classroom. The evidence collected in this work suggests that translingual assessments are worth more exploration in Mbandaka and contexts that face similar challenges. We encourage test developers to explore more rich forms of translanguaging in mathematics assessments. When doing so, it is important to approach the task without preconceptions about language use and to develop tests that follow the linguistic patterns of the population being assessed. Ultimately, the purpose of translanguaging in assessment is to honor the ways in which students speak language (Swain 2006), not to impose exogenous ways of flexible bilingualism or multilingualism.

Translanguaging has not received enough attention and support from practitioners in international education. To transform the practice of mathematics assessments we need to observe a political will to shift from a monolingual to a multilingual approach to assessments, which is not granted. Multilingual approaches are not settled by only including bilingual students in the target and pilot populations. Implementing multilingual approaches requires a bolder commitment to dissolve the complex and relatively unquestioned monolingual language assumptions that pervade measurement

and evaluation practice. Moving toward a translingual framework in mathematics assessments requires reconstructing our notions about language and educating test users. Translanguaging can improve the opportunities of multilinguals to engage in mathematics assessments and demonstrate what they know and can do, in ways that recognize their language practices and beliefs, and in ways that do not attack their identities or senses of place. The benefits of translanguaging can only flourish with a strong and intentional commitment on the side of researchers and practitioners. We hope that this study helps move the conversation forward.

References

Abedi, J. 2004. "The No Child Left Behind Act and English Language Learners: Assessment and Accountability Issues." *Educational Researcher* 33 (1): 4–14.

Abedi, J. 2009. "Computer Testing as a Form of Accommodation for English Language Learners." *Educational Assessment* 14 (3/4): 195–211.

Abedi, J., and P. Gandara. 2006. "Performance of English Language Learners as a Subgroup in Large-Scale Assessment: Interaction of Research and Policy." *Educational Measurement: Issues and Practice* 25 (4): 36–46.

Adler, J. 2001. *Teaching Mathematics in Multilingual Classrooms*. Dordrecht: Kluwer.

Adoniu, M., and Y. Quing. 2014. "Language, Mathematics, and English Language Learners." *Australian Mathematics Teacher* 70 (3): 3–13.

Ascenzi-Moreno, L. 2018. "Translanguaging and Responsive Assessment Adaptations: Emergent Bilingual Readers through the Lens of Possibility." *Language Arts* 95 (6): 355–69.

Bailey, A. L. 2007. *The Language Demands of Schools: Putting Academic English to the Test*. New Haven, CT: Yale University Press.

Baker, C. 2011. *Foundations of Bilingual Education and Bilingualism*. Bristol: Multilingual Matters.

Barwell, R., L. Chapsam, T. Nkambule, and M. Setati-Phakeng. 2016. "Tensions in Teaching Mathematics in Contexts of Language Diversity." In *Mathematics Education and Language Diversity*, ed. Richard Barwell, Philip Clarkson, Anjum Halai, Mercy Kazima, Judit Moschkovich, Nuria Planas, Mamokgethi Setati-Phakeng, Paola Valero, and Martha Villavicencio Ubillus. Cham: Springer.

Barwell, R., and D. Pimm. 2016. "Bakhtin and Some Dilemmas of Mathematics-Language." In *Mathematics Education in a Context of Inequity, Poverty and Language Diversity*, ed. M. Phakeng and S. Lerman. Cham: Springer.

Branson, J., and D. Miller. 2007. "Beyond 'Language': Linguistic Imperialism, Sign Languages and Linguistic Anthropology." In *Disinventing and Reconstituting Languages*, ed. S. Makoni and A. Pennycock. Clevedon: Multilingual Matters.

Brown, T. 2002. *Mathematics Education and Language: Interpreting Hermeneutics and Post-Structuralism*. Boston: Kluwer.

Busch, B., and J. Schick. 2007. "Educational Materials Reflecting Heteroglossia: Disinventing Ethnolinguistic differences in Bosnia-Herzegovina." In *Disinventing and Reconstituting Languages*, ed. S. Makoni and A. Pennycook. Clevedon: Multilingual Matters.

Canagarajah, Suresh. 2011. "Codemeshing in Academic Writing: Identifying Teachable Strategies of Translanguaging." *Modern Language Journal* 95 (3): 401–17.

Creese, A., and A. Blackledge. 2010. "Translanguaging in the Bilingual Classroom: A Pedagogy for Learning and Teaching?" *Modern Language Journal* 94 (1): 103–15.

Cummins, J. 2000. *Language, Power and Pedagogy: Bilingual Children in the Crossfire*. Clevedon: Multilingual Matters.

DiCerbo, P. A., K. Anstrom, L. Baker, and C. Rivera. 2014. "A Review of the Literature on Teaching Academic English to English Language Learners." *Review of Educational Research* 84 (3): 446–82.

Escamilla, K. 2006. "Monolingual Assessment and Emerging Bilinguals: A Case Study in the US." In *Imagining Multilingual Schools*, ed. Ofelia Garcia, Tove Skutnabb-Kangas and Maria E. Torres-Guzman. New York: Multilingual Matters.

García, O. 2009. *Bilingual Education in the 21st Century: A Global Perspective*. Malden, MA: Wiley.

García, O., S. Ibarra Johnson, and K. Seltzer. 2017. *The Translanguaging Classroom: Leveraging Student Bilingualism for Learning*. Philadelphia: Caslon.

Garcia, O., J. Kleifgen, and L. Falchi. 2008. *From English Language Learners to Emergent Bilinguals*. New York: Campaign for Educational Equity.

Garcia, O., T. Skutnabb-Kangas, and M. Torres-Guzman. 2006. "Weaving Spaces and (de)constructing Ways for Multilingual Schools: The Actual and the Imagined." In *Imagining Multilingual Schools: Languages in Education and Glo(c)alization*, ed. Ofelia Garcia, Tove Skutnabb-Kangas and Maria E. Torres-Guzman. Clevedon: Multilingual Matters.

Gee, J. 1999. *An Introduction to Discourse Analysis: Theory and Method*. London: Routledge.

Global Education Monitoring Report. 2016. *If You Don't Understand, How Can You Learn?* Paris: GEM Report.

Goodman, J. 2013. "The Meritocracy Myth: National Exams and the Depoliticization of Thai Education." *Journal of Social Issues in Southeast Asia* 28 (1): 101–31.

López, A., S. Turkan, and D. Guzman-Orth. 2017. *Conceptualizing the Use of Translanguaging in Initial Content Assessments for Newly Arrived Emergent Bilingual Students*. Princeton, NJ: Educational Testing Service.

Makoni, S., and P. Mashiri. 2007. "Critical Historiography: Does Language Planning in Africa Need a Construct of Language as Part of Its Theoretical Apparatus?" In *Disinventing and Reconstituting Languages*, ed. S. Makoni and A. Pennycook. Clevedon: Multilingual Matters.

Makoni, S., and A. Pennycook. 2007. "Disinventing and Reconstituing Languages." In *Disinventing and Reconstituting Languages*, ed. S. Makoni and A. Pennycook. Clevedon: Multilingual Matters.

Martiniello, M. 2008. "Language and the Performance of English-Language Learners in Math Word Problems." *Harvard Education Review* 78 (2): 333–68.

National Governors Association. 2012. *Common Core Standards State Initiative: Mathematics Standards*. Washington, DC: National Governors Association. http://www.corestandards.org/Math/.

Platas, L., L. Ketterlin-Gellar, L. Brombacher, and Y. Stiabkhan. 2014. *Early Grade Mathematics Assessment (EGMA) Toolkit*. Research Triangle Park, NC: RTI International.

Popham, W. J. 2001. *The Truth about Testing: An Educator's Call to Action*. Alexandria, VA: Association for Supervision and Curriculum Development.

Popham, W. J. 2003. "Trouble with Testing: Why Standards-Based Assessment Doesn't Measure Up." *American School Board Journal* 190 (2): 14–17.

Popham, W. J. 2008. "When the Test Says You're Not So Smart." *Educational Leadership* 65 (7): 87–88.

Popham, W. J. 2016. "Standardized Tests: Purpose Is the Point." *Educational Leadership* 73 (7): 44–49.

Reubens, A. 2009. *Early Grade Mathematics Assessment (Egma): A Conceptual Framework Based on Mathematics Skills Development in Children*. Research Triangle Park, NC: RTI International. https://www.eddataglobal.org/documents/index.cfm?fuseaction=pubDetail&id=193.

Rotberg, I. 1995. "Myths about Test Score Comparisons." *Science* 270 (5241): 1446–48.

Setati, M. 2008. "Access to Mathematics versus Access to the Language of Power: The Struggle in Multilingual Mathematics Classrooms." *South African Journal of Education* 28 (1): 103–16.

Shohamy, E. 2011. "Assessing Multilingual Competencies: Adopting Construct Valid Assessment Policies." *Modern Language Journal* 95 (3): 418–29.

Shohamy, E. 2006. "Imagined Multilingual Schools: How Come We Don't Deliver." In *Imagining Multilingual Schools*, ed. Ofelia Garcia, Tove Skutnabb-Kangas, and Maria E. Torres-Guzman. New York: Multilingual Matters.

Sireci, S., and F. Gandara. 2016. "Testing in Educational and Developmental Settings." In *The ITC International Handbook of Testing and Assessment*, ed. F. T. L. Leong, D. Bartram, and F. Cheung. New York: Oxford University Press.

Swain, M. 2006. "Languaging, Agency, and Collaboration in Advanced Language Proficiency." In *Advanced Language Learning: The Contribution of Halliday and Vygotsky*, ed. H. Byrnes. London: Continuum.

Thurlow, M., and R. Kopriva. 2015. "Advancing Accessibility and Accommodations in Content Assessments for Students with Disabilities and English Learners." *Review of Research in Education* 39 (1): 331–69.

Tollefson, J. W. 2008. "Language Planning in Education." In *Language Policy and Political Issues in Education*, ed. Nancy H. Hornberger. New York: Springer.

UNESCO. 2016. *Global Education Monitoring Report 2016. Education for People and the Planet: Creating Sustainable Futures for All*. Paris: UNESCO.

UNESCO. 2017. *Global Education Monitoring Report 2017/8. Accountability in Education: Meeting our Commitments*. Paris: UNESCO.

Velasco, P., and O. Garcia. 2014. "Translanguaging and the Writing of Bilingual Learners." *Bilingual Research Journal: The Journal of the National Association for Bilingual Education* 37 (1): 6–23.

Wells, A. 2014. "Seeing Past the 'Color-Blind' Myth of Education Policy: Addressing Racial and Ethnic Inequality and Supporting Culturally Diverse Schools." *Education Digest* 80 (3): 38–41.