

Are WIDA Test Results Appropriately Reflecting Multilingual Learners' Language Skills According to ESOL Teachers' Experiences? Results of a Pilot Study

GATESOL Journal 2024,
Vol. 33(1), pp. 3–17
doi:10.52242/gatesol.184
ISSN: 2331-6845

Emily Patterson
Elke Schneider

Richard W. Riley College of Education, Winthrop University

Abstract

Within the field of multilingual learner (ML) education, ESOL teachers' voices are often overlooked and underrepresented despite their integral role in developing productive and knowledgeable future citizens. This study sheds light on the experiences of ESOL teachers that administer the federally mandated annual standardized testing created by World-Class Instructional Design and Assessment (WIDA). It uses focus-group interviews to gather qualitative data from two neighboring school districts in order to gain insight into the WIDA assessments' reflection of language proficiency. WIDA testing evaluates English language development by measuring academic and social language skills within the four language domains: reading, writing, listening, and speaking. MLs without severe, classified disabilities in grades K-12 participate annually in the Assessing Comprehension and Communication of English State-to-State (ACCESS) assessment by WIDA. Results place students into proficiency categories ranging from entering (a student new to English) to bridging (a near-fluent ML), as defined by the WIDA language proficiency standards. The interviews allowed the ESOL teachers to share their experiences with MLs' classroom performances in the four testing domains compared to their performances on the online WIDA assessments. The interview data was analyzed and categorized into three main themes based on teacher responses: assessment preparation techniques, common perceptions of the assessment, and proposed changes to improve the assessment. Each of these themes, with their local and national implications, are discussed as they affect the nature of ESOL instruction and assessment.

Keywords

language proficiency assessment, multilingual learners, WIDA ACCESS, ESOL teacher perspectives

Introduction

Appropriate instruction and assessment of multilingual learners (MLs) within education has become an increasingly important field of study. As the percentage of MLs within public schools continues to rise, research within the field of ESOL, especially pertaining to the assessment of

Emily Patterson, Student, Winthrop University, Rock Hill, SC; email patterson9@winthropalumni.com.

Elke Schneider, Professor, Winthrop University, Rock Hill, SC; email schneidere@winthrop.edu

English language development, remains limited. The purpose of this study is to explore the experiences of ESOL teachers from two local school districts in the east coast region of the U.S. regarding the nature and authenticity of WIDA ACCESS testing (Wisconsin Center for Education Research, 2023). WIDA ACCESS assessments provide data that indicate the development of English in students whose first language is not English to appropriately guide ensuing instruction. ESOL teachers share personal experiences through focus group interviews to shed light on the various perspectives of educators regarding the authenticity of the WIDA ACCESS test results in measuring English language development. Authenticity, in the context of this study, is defined by Varga and Guignon (2020) as “a reliable, accurate representation” of MLs’ English language proficiency (p. 1).

This study is significant due to general ESOL teachers’ perceptions of, and anecdotal references to, the discrepancies between student performance on standardized WIDA tests and perceived English language development within the classroom. These anecdotal comments have alluded to challenges with WIDA testing as an adequate representation of MLs’ language development based on teachers’ daily interactions and continual monitoring of language acquisition in the classroom. These teachers share the seemingly common frustrations of perceiving the WIDA language proficiency assessments as an ineffective measure of language proficiency across grade levels, schools, and districts. Although most U.S. states measure English language proficiency using WIDA, limited independent research exists exploring teachers’ perceptions of the assessments as an accurate measure of language development (King & Bigelow, 2018). Currently, the existing body of independent research into the authenticity of WIDA language proficiency assessments is limited to the quantitative findings of Coulter (2017) and the qualitative findings of Lopez and Garcia (2020) and Waters (2020). Coulter (2017) compared WIDA ACCESS proficiency ratings with previous alternative standardized testing. Lopez and Garcia (2020) interviewed ESOL teachers and found out that most teachers struggle to utilize standardized data for instruction. Moreover, Waters (2020) examined the variance in teacher perspectives of WIDA ACCESS testing based on demographics and teaching environments to determine the impact of test results on ESOL policy and decision-making. These findings support the inquiry into the authenticity of WIDA ACCESS testing.

To the knowledge of the authors to this date, ESOL teachers’ experiences with the preparation for, administration of, and comparison of test results with in-class language performances of MLs have not been explored in the ESOL literature. Therefore, this study creates a unique forum to share teachers’ experiences through inquiry into pedagogical techniques and beliefs surrounding the standardized assessments. Moreover, the consistently growing number of MLs in public schools creates an increasing demand for appropriate and authentic testing to track language development (Ariza & Coady, 2018).

The following sections provide readers with background information to better understand the relevance of this study in terms of (a) MLs’ language acquisition, (b) legal support for appropriate language performance assessment, and (c) the nature of WIDA ACCESS testing.

MLs and Language Acquisition

When acquiring a second or additional language, MLs develop language in two domains: basic interpersonal communication skills (BICS) and cognitive academic language proficiency (CALP) (Cummins, 2016). Developed and defined by Cummins (2016), BICS encompasses oral conversational fluency that generally develops within one to three years of consistent social immersion in an additional language and culture. CALP includes spoken or written academic

language used in informational exchanges, typically within a formal academic setting. On average, MLs develop near-fluent CALP over the span of five to seven years with regular exposure in an academic setting and up to nine years with limited exposure (Cummins, 2008). However, Cummins' theory of language acquisition receives criticism for "creating a binary view of language" that does not always adequately reflect language acquisition (Scarcella, 2003, p. 6). In her critique, Scarcella (2003) refers to Cummins' model as a stagnant representation of language development encompassing fixed choices of language comprehension and production. However, she overlooks the progressivity of Cummins' model, in which students continuously acquire new social and academic language skills simultaneously in diverse environments (Ariza & Coady, 2018). Cummins' theory of language acquisition remains pertinent to the assessment of language proficiency because, when measuring language development, language proficiency assessments should consider the disparity between social and academic language acquisition.

MLs' language acquisition is influenced by the learning environment (Makarova, 2019), where they can gain exposure to and authentic experiences with using social language to engage in social interactions that serve as the dominant language function within a positive setting. However, authentic experiences using academic vocabulary can be limited because the language is often decontextualized and used primarily in academic endeavors (Ariza & Coady, 2018). This decontextualization results from the expectation that students utilize academic vocabulary solely within content-specific settings, such as language unique to geometry during mathematics instruction, rather than integrating domain-specific vocabulary in a variety of authentic settings. Furthermore, language-level properties and child-level properties (Paradis, 2019) can affect MLs' language acquisition rates. Language-level properties include features that impact language acquisition, such as the complexity and structural dynamics of the new language. Child-level properties are properties that relate to the individual learner to influence language acquisition, including input and output opportunities and cognitive capacity (Coulter, 2017; Paradis, 2019). Within these domains, MLs possess complex, higher-order thinking and executive functioning skills but struggle to verbalize their thoughts in a new language, resulting in a language gap (Coulter, 2017, p. 1) rather than an achievement gap (Kudo & Swanson, 2014). The educational environment also contributes to the language gap by limiting input and output opportunities on which language acquisition heavily relies. For example, classroom instruction that limits peer discussions prevents MLs from developing academic oral language during critical thinking tasks. The properties of language acquisition combined with knowledge of the complex capabilities of MLs create a foundation for effective ESOL pedagogy and assessment.

Legal Support for Appropriate Performance Assessment

The legal support for appropriate language proficiency tests in public schools is based on several legal developments starting in the early 1960s. These landmark court cases related to ML populations and ESOL education, including *Lau v. Nichols* (1974), *Plyler v. Doe* (1982), *Castañeda v. Pickard* (1981), and *United States v. Texas* (1970) (McEachern, 2022), impact the pedagogy and classroom experiences of MLs. The first federal acknowledgement of the needs of MLs within public education deemed segregation of schools unlawful and mandated federal grants to support programs for MLs (Bilingual Education Act, 1968). As ML populations rapidly expanded, claims arose that identical education did not present equal educational opportunities for MLs. The case ruled that schools must take affirmative actions to remove educational barriers for all students, including MLs (*Lau v. Nichols*, 1974). Subsequently, multiple court cases led to ESOL programs in public education by requiring education for undocumented immigrants (*Plyler*

v. Doe, 1982), research-based pedagogy (Castañeda v. Pickard, 1981), and progress monitoring through standardized testing (United States v. Texas, 1970). See Appendix A for a detailed overview of these court cases.

Nature of WIDA ACCESS Testing

MLs are federally mandated to complete a large-scale, norm-referenced, summative language proficiency assessment on an annual basis, as required by Castañeda v. Pickard (1978). WIDA currently provides assessments to 41 U.S. states and 2 U.S. territories. Originally funded by a U.S. Department of Education grant, WIDA was established in 2003 by the Wisconsin Department of Public Instruction to design language proficiency standards and assessments for MLs (Wisconsin Center for Education Research, 2023). To identify MLs' home languages, new students in U.S. public education must complete a home language survey (Castañeda v. Pickard, 1978). MLs without classified disabilities in grades 1-12 take the online ACCESS test administered by a local ESOL teacher and scored by affiliates of the WIDA organization. The WIDA ACCESS test determines MLs' English proficiency based on vocabulary usage, language forms and conventions, and linguistic complexity in the following content areas: social and instructional language, English language arts, mathematics, science, and social studies (Wisconsin Center for Education Research, 2023). According to WIDA-based research, the ACCESS assessment aligns national Common Core State Standards (CCSS) with WIDA language standards (Cook, 2014; Wisconsin Center for Education Research, 2023) by assessing content knowledge and critical thinking skills (Coulter, 2017).

Students take this online test in four separate parts for a total of 265 testing minutes. Local ESOL teachers administer and supervise the test, following an administration model provided by WIDA. Assessed language domains include reading, writing, listening, and speaking (see Appendix B1). ESOL teachers then use the test scores to guide future instruction based on MLs' strengths and areas for improvement. WIDA uses criterion-referenced scoring to place students at different levels of language proficiency, as described in Appendix B2 (Coulter, 2017). The six levels of language acquisition are associated with a number system (1-6) starting at entering (1) for students with minimal language proficiency, beginning (2), developing (3), expanding (4), bridging (5), and reaching (6) (Wisconsin Center for Education Research, 2023).

The online nature of this test impacts MLs' language performances in a variety of ways. According to Coulter (2017), teachers consider the negative effects of extended online testing on MLs' performances when reviewing test scores. Also, having to record themselves speaking into a computer using a headset and microphone creates an unnatural language interaction for MLs that differs from speaking with others in person (Coulter, 2017). Additionally, the writing portion requires students to type constructed responses within a time limit and deviates from MLs' writing tasks with a writing utensil and paper. Coulter's (2017) WIDA-independent study also describes the disconnect in test scores and classroom performances because the assessment requires MLs to apply grade-level content knowledge and critical thinking skills in English without considering MLs' deficits in both areas. According to Coulter (2017), only 20 percent of the testing material measures social language skills, which is evident in the updated WIDA English Language Standards Statements (2020). Only Standard 1 explicitly addresses the assessment of "language for social and instructional purposes" (Wisconsin Center for Education Research, 2023) while the remaining four language standards for language arts, mathematics, science, and social studies exclusively emphasize academic language and content knowledge (WIDA, 2020, pp. 24-25; Wisconsin Center for Education Research, 2023).

Research Methods

This study delved into the WIDA ACCESS tests used to measure language proficiency to answer the research question: How do ESOL teachers perceive the WIDA ACCESS language proficiency assessment as an authentic representation of MLs' English language proficiency? Discussions in focus-group interviews were chosen for data collection in order to allow participants to interact with each other and share their personal experiences with WIDA testing.

Nine of 34 invited ESOL teachers from two local suburban school districts voluntarily participated in virtual focus group interviews. Participants were identified using a sample of convenience by locating the names of local ESOL teachers on district websites and contacting them directly through email to request participation. Following university International Review Board (IRB) protocols, participation was voluntary, and participant identities remained anonymous. All participants were female, white, native speakers of English who had previous or current experiences as elementary ESOL teachers in two contiguous school districts. Three of the nine participants were employed in the largest school district in the county of interest, and, at the time of the study, 16.5% of the student population were classified as MLs according to the local school district website. The remaining six participants were employed in an adjacent school district with similar demographics. Three participants taught in both middle school and elementary ESOL classrooms but shared their experiences specifically pertaining to the administration of WIDA assessments in elementary settings to maintain consistency and specificity in the gathered data. Four participants requested individual interviews. One session consisted of two interviewees, and another session held three participants.

During 30-minute individual and group interviews, six prepared questions were posed as follows. The researchers clarified any necessary wording, such as defining authenticity in the context of this study, to ensure participants fully understood each question and answered the questions to the best of their abilities.

1. In what ways do you prepare your students for WIDA ACCESS testing, and can you provide examples of preparation techniques?
2. Which of the four domains (reading, writing, listening, and speaking) is harder or easier to prepare students for? Why?
3. In your personal experience as an ESOL teacher, how has the online nature of the assessment affected student performance when compared to your informal assessments and interactions with the students?
4. To what degree does the WIDA test reflect MLs' English proficiency in the four testing domains?
5. What changes could be made to the WIDA assessment to ensure students demonstrate their English proficiency with more authenticity than the current assessment results?
6. Based on your experiences, to what degree does WIDA testing require students to access cultural understandings, prior academic knowledge, and keyboarding skills?

Findings

The researchers analyzed the interview transcripts generated automatically by the virtual meeting platform. They independently coded the transcripts according to recurring themes, discussed commonalities in their theme-coding, and resolved coding differences through negotiated agreements (Belotto, 2018), ensuring interrater reliability. They then used the coded transcripts to

identify three common themes within the data: (a) teacher preparation techniques for the test, (b) teacher perceptions of the test, and (c) proposed changes to the test.

Theme 1: Preparation Techniques

The first theme, preparation techniques for the test, was separated into four categories: preparation for speaking, writing, reading, and listening. All interviewees commented on common preparations, such as conferencing with students to discuss the previous year's test scores (7 out of the 9 participants), setting goals for the upcoming year (6/9), organizing students into ability-based small groups (2/9), and practicing sample test questions (5/9) to "simulate the test as much as possible," as one teacher stated.

Speaking test preparations. The interviewees built a comprehensive list of preparation techniques to improve each skill required for a satisfactory test score. For instance, eight of the nine teachers discussed incorporating recording technology to allow students to practice physically speaking into a microphone with a headset as a means to simulate the test format. Five teachers mentioned practicing the mechanical functions of clicking the record button, speaking for a set time, and then ending the recording. To familiarize students with talking about a topic for an extended period, one teacher has students describe a displayed picture using details and academic vocabulary while timing responses with a running sand timer. All nine interviewees emphasized the importance of acclimating students to the test format rather than providing direct instruction to improve MLs' oral responses. Lastly, four teachers have students identify relevant academic and content-specific vocabulary to track the vocabulary usage in their responses so students "use as much academic language as possible."

Writing test preparations. Similar to the speaking preparation, seven teachers simulate the writing portion of the test in their preparation. To mirror test formats, students in grades 1-3 write responses on paper while students in grades 4-5 type responses on the computer. Three teachers discussed their focus on developing the writing process with brainstorming, drafting, revising, and editing responses. These same three teachers also dedicate units of study to each of the writing functions, such as comparing and contrasting or cause-and-effect. One teacher uses a color-coding strategy to ensure students effectively organize their responses. Three participants also discussed a technique similar to the speaking preparation in which students describe a picture using details and academic vocabulary. The majority of described writing preparation emphasizes identifying and incorporating key vocabulary and academic language into the responses. Four teachers also push into general education classrooms with their MLs in a separate academic environment to connect academic content to their writing.

Reading test preparations. Similar to speaking and writing, a common preparation technique used to familiarize MLs involves simulating the test format by using practice questions that require students to read and answer content-related questions. Four teachers found this crucial because they noticed MLs getting confused by unfamiliar academic content when answering the questions. For example, questions may provide students with a brief passage involving mathematics and U.S. currency. Although the answer to the question is provided in the given passage, MLs can become distracted by the mathematical nature of the question as well as a possible lack of prior knowledge surrounding U.S. monetary values. This distraction often affects MLs' abilities to apply language skills to effectively answer the question. Therefore, eight of the

nine teachers pose discipline-specific questions repeatedly. Additionally, these eight teachers mentioned that the instruction provided within general education classrooms, combined with their direct support, strengthened MLs' reading skills. Thus, all interviewees did not see a need to emphasize direct preparation for the reading portion of the test.

Listening test preparations. All nine interviewees direct little attention to listening practices because both social and academic listening skills develop naturally within school environments, which provides MLs with daily listening practice in their classrooms. One teacher described having MLs listen to their own recordings from the speaking practice to discuss strengths and areas for improvement using student-friendly rubrics provided by WIDA. Another teacher records prompts for students to listen to and answer questions that mirror the test. Overall, the interviewed ESOL teachers prepare MLs for each of the four testing domains similarly by mirroring the test format and emphasizing the development of academic language.

Theme 2: Teacher Perceptions of the Assessment

The second theme encompasses teachers' perceptions of the test and its ability to appropriately measure English proficiency. Two main categories arose from these perceptions: domain-specific difficulties and emphasis on academic content.

Domain-specific difficulties. All nine participants agreed that the listening portion of the test poses the least difficulty for MLs, and they devote minimal direct instructional time to listening test preparations. One teacher credited the lack of difficulty to the natural academic and social listening practice occurring in general education classrooms. Six teachers observed consistently higher scores on the listening and reading domains' multiple-choice format and partially attributed MLs' success to the limited pool of answer choices and frequent exposure to multiple-choice questions. Also, seven participants attributed annual growth in reading scores to adequate practice in the general education classroom. Five interviewees consistently identified the reading portion as more difficult than the listening portion due to content-specific questions. They noticed that some MLs tend to use academic background knowledge rather than language skills to answer content-based questions. To illustrate, and referring to the aforementioned example of students encountering a question based on U.S. currency and mathematics, interviewees described MLs' tendency to use knowledge of U.S. currency to answer the question rather than their reading skills to locate the answer provided in the passage. Despite these difficulties, all nine teachers found the reading and listening portions less difficult than the speaking and writing domains.

In contrast, every teacher identified speaking and writing test sections as significantly more difficult than listening and reading tasks. All nine teachers attributed the difficulty of the speaking test first to the unnatural test format and repeatedly described MLs' struggles with the process of recording their answers while speaking to a screen rather than a real person. They also pointed out that recording an oral response for an extended period without verbal or nonverbal feedback (i.e., prompting or nodding) creates, according to one teacher, an "intimidating" setting. Another teacher shared, "I have kids who speak perfect English who score very, very low on [speaking]" because of the students' discomfort with the response format. Additionally, eight of the nine teachers mentioned that the speaking test sets higher expectations for language than commonly found in daily conversation by requiring students to use academic vocabulary not typically used in everyday speech. To combat this discrepancy, during class time two teachers rely heavily on having MLs speak in complete sentences using relevant academic vocabulary. All nine teachers described the writing portion of the test as comparably difficult to the speaking test because the writing format

differs from state-standardized tests in which students conduct prompted text-dependent analyses with an expected response length (Writing Component, 2022). Seven interviewees also pointed out that the writing tasks require writing stamina beyond developmentally appropriate levels. MLs respond to three prompts that each demand lengthy responses with academic language and relevant content. All teachers commented that questions often referenced academic content, and four participants found that the topics were often unfamiliar or unnatural to write about for MLs. In sum, teachers perceived the writing and speaking portions to be the most difficult because of their unnatural format and the high expectation of students' command of academic language.

Emphasis on academic content. All nine participants commented on the test's emphasis on academic language and content knowledge. Six teachers viewed the content-oriented questions as inaccurate measures of language proficiency because "language proficiencies come out much lower than where the child actually is," as one teacher stated. Students receive daily scaffolding and academic support, such as visual aids, but the test provides minimal support when students use academic vocabulary. Teachers perceived the test as "very content-oriented" and "testing [MLs'] academic language" rather than measuring language proficiency. Five interviewees also noted that one challenge for MLs was that relevant grade-level standards have often not been taught at the time of testing. These circumstances create a disconnect between the test's expected prior knowledge and the students' actual knowledge, despite attempts by WIDA to align the assessment to national Common Core standards (Center for Public Education, 2014). One teacher commented, "Trying to get a student who already struggles with the language to write about something that they don't have any background knowledge about makes zero sense." Overall, the teachers perceived the test as content-driven rather than focused on language.

Theme 3: Proposed Changes

The final theme presents proposed changes to improve the testing experience for MLs and teachers. Suggested changes include: (a) comparison of MLs' test results with those of native English speakers, (b) alterations to the test format, (c) consideration of multiple data points to measure language proficiency, and (d) administration of the test later in the school year.

Comparing test results of MLs and native speakers of English. Each interviewee stressed the importance of defending the validity of the test by having native English speakers take an identical test to compare their results with those of the MLs and provide insights into where the test fails to accurately measure language proficiency. Eight teachers expressed their doubts that a native speaker would receive a passing score because of the length, format, and content of the test. One teacher cited an instance in which an ESOL teacher appealed to WIDA by asking for a native speaker to take the test, but WIDA representatives stated the results would be insufficient to disprove the accuracy of the test because of the discrepancy in language abilities between native speakers and MLs. Therefore, one teacher asked, "If a native English speaker can't pass this test, how do you expect this [ML] to?" All nine interviewees agreed that comparing the test scores of native speakers and MLs could improve performance expectations.

Altering test format. Interviewees provided multiple suggestions to maximize the authenticity of the test results, such as making the format developmentally appropriate by balancing the online and written portions of the test. While the kindergarten version of the ACCESS test is taken almost entirely on paper and involves developmentally appropriate skills, such as allowing for the use of manipulatives, the first-grade version of the test requires students

to complete three of the four testing domains online. This format requires developmentally inappropriate technology skills and testing stamina, as exemplified by one teacher who stated, “I hardly ever have any first-grade students that exit the ESOL program.” Existing research by Waters (2020) on teacher perceptions of the WIDA ACCESS assessment also supports the use of developmentally appropriate testing practices by arguing for a shift back to paper forms to center the focus on language proficiency rather than computer and keyboarding skills. Moreover, every teacher discussed the unnatural format of recording the speaking portion on a computer. Six of these teachers mentioned the benefits of administering the speaking test in person to create a more natural testing environment for MLs. Six participants described students tiring during the test and rushing as their concentration fades near the end, particularly in the constructed-response portions of the test, including speaking and writing. Overall, these teachers felt that by reducing the testing time spent online, young MLs were less likely to exceed their testing stamina.

The final suggestion involves content-reliant questions. Five teachers proposed including more social language, and three of these teachers recommended utilizing academic content from the previous grade levels’ standards rather than those of the current grade. This would decrease students’ encounters with unfamiliar content and allow students to use higher-order thinking processes while measuring students’ daily language capabilities.

Using multiple data points to determine language skills. All nine teachers recommended collecting and triangulating data to determine language proficiency with summative and formative assessments to “look at a child holistically,” as one teacher described. The court case *United States v. Texas* (1970) also supports the use of additional data points to measure language proficiency development by demanding appropriate progress monitoring of MLs’ language development. Unanimously, participants perceived the test as an inaccurate measure of language proficiency. The most common suggestions for improved measures of proficiency included using student portfolios and running records (7/9) and gathering various formative assessments throughout the school year (6/9).

Altering ACCESS testing time. Four teachers proposed a later test administration time to a) provide teachers with more instructional time before assessing student growth, and b) provide students with more time to improve their language and content proficiencies. Currently, testing in this district occurs in February, leaving nearly three months of further instruction and possible proficiency growth unassessed. Additionally, ESOL teachers first gather data on MLs at the start of the school year and then begin working with them in early October, leaving students with only three to four months of instruction from ESOL teachers before taking the test. While four teachers also considered possible conflicts with annual state standardized testing taken near the end of the school year, they continued to stress the shortcomings of the current testing dates.

Implications

Despite the small scale of this study, many significant local and national implications became evident that are relevant to the ESOL field as a whole. Locally, one crucial implication arose based on the teachers’ perceptions of the varying difficulty of the test through consistent anecdotal comments about students feeling “stuck” in the program as student test scores did not align with their perceived language proficiency. A feasible solution to remedy this identified disconnect between test performance and perceived language proficiency would be to alter the state exit criteria. The state’s ESOL exit criteria is based solely on student performance on the WIDA ACCESS test, as students must receive a 4.4 composite score and 4.0 in each individual testing

domain (WIDA, 2023). This single test determines whether students “are just going to be MLs for life because they can’t pass the test,” as one teacher noted, despite appearing proficient in daily activities. Lowering the individual domain requirements while maintaining the composite score would relieve testing pressures to create a more authentic, and holistic, reflection of students’ abilities.

Nationally, alterations to the exit criteria would also improve the authenticity of language proficiency measures. One change includes considering multiple data points to triangulate data, make educational decisions, and assess progress to allow ESOL educators to determine MLs’ language proficiency more accurately. Educators rely on the most effective assessment practices to measure and reflect language development and proficiency. A single test score, however, appears to not reflect MLs’ authentic language capabilities based on this study’s findings. Effective assessment methods integrate natural language measures and provide multiple opportunities for authentic expressions of language proficiency (Coulter, 2017). This study contributes to Coulter’s research on WIDA testing by providing a unique perspective with qualitative interview data from in-service ESOL teachers and further supports the use of multiple data points to measure language proficiency. Secondly, the WIDA Consortium could benefit from accepting input from ESOL teachers with experience implementing ACCESS testing and a multifaceted knowledge of their MLs. Valuable suggestions to consider include aligning testing situations with the classroom environment, integrating more social language while reducing the number of content-related questions, and providing testing later in the year so MLs can demonstrate growth with more authenticity.

These implications provide insight into the challenges of the WIDA ACCESS test and possible solutions to improve the authenticity of the assessment. Moreover, these findings could be considered to improve aspects of standardized assessments within general and specialized educational fields. For example, holistic assessment approaches such as collecting multiple data points improve the quality of instruction and assessment throughout the field of education. Therefore, this research and its findings serve to benefit the entire educational community.

Conclusion

This study intended to shed light on elementary ESOL teachers’ experiences with WIDA ACCESS testing regarding its authenticity compared to MLs’ proficiency in the classroom. The passion with which the teachers shared their views encourages future research on this topic. In response to the drastically increasing number of MLs in the U.S. (Park et. al., 2017), further studies could help improve teacher and student experiences with WIDA ACCESS testing and the quality of language assessments.

Our findings contribute to a small body of research related to WIDA ACCESS testing. While WIDA-affiliated studies identify WIDA ACCESS testing as effective for decision-making based on standardized test data, pedagogical approaches, and evaluation of teachers, independent studies like ours identify several challenges with the efficacy of WIDA ACCESS testing. These are validated by Coulter (2017), Waters (2020), and Lopez and Garcia (2020). These independent studies highlight the need for improved standardized language proficiency measures to combat the discrepancy between measured and perceived language proficiencies of MLs (Coulter, 2017; Lopez & Garcia, 2020; Waters, 2020). However, WIDA ACCESS tests currently serve as the sole consideration to determine ML’s language proficiency.

Further studies are needed to appraise the authenticity of WIDA ACCESS testing from the perspective of ESOL teachers to expand upon this small-scale study. Such studies could include

geographic, socioeconomic, and ethnographic diversity with larger participation pools, including voices of early childhood, elementary, middle school, and high school ESOL experiences along with responses from rural, suburban, and urban schools. Additionally, it would be insightful to compare WIDA assessment practices and exit percentages with results from other standardized language testing used in states that do not implement WIDA ACCESS testing. Overall, the ESOL field and the educational community can benefit from the findings of this study and bring the improvement of ESOL education and assessment to the forefront of educational research.

References

- Acheson, K., & Gall, M. D. (1997). *Techniques in the clinical supervision of teachers: Preservice and inservice applications* (4th ed.). Longman.
- Ariza, E., & Coady, M. (2018). *Why TESOL? Theories & issues in teaching English to speakers of other languages in K–12 classrooms* (5th ed.). Kendall Hunt.
- Belotto, M. (2018). Qualitative research: Managing the challenges of coding, interrater reliability and thematic analysis. *Qualitative Report*, 23(11), 2620–2633.
- Bilingual Education Act, 90 US 247 (1968). <https://www.govinfo.gov/content/pkg/STATUTE-81/pdf/STATUTE-81-Pg783.pdf#page=32>.
- Castañeda v Pickard, 648 F.2d 989 (1978). <https://web.stanford.edu/~hakuta/www/LAUcIAPolicy/IA1bCastanedaFullText.htm>
- Center for Public Education (2014). *Understanding the common core*. Center for Public Education.
- Cook, G. H. (2014). *Examining relationships between alternative ACCESS and state alternative assessments: Exploring the notions of English language proficiency*. University of Wisconsin. <https://wida.wisc.edu/sites/default/files/resource/Report-Examining-Relationships-Between-Alternate-Access-and-State-Alternate-Assessments-AL.pdf>
- Coulter, K. (2016). *Effectiveness of WIDA English language development standards meeting proficiency for English language learners* [Doctoral dissertation, Tennessee State University]. ProQuest.
- Cummins, J. (2008). *BICS and CALP: Empirical and theoretical status of the distinction*. In N.H. Hornberger (Ed.), *Encyclopedia of Language and Education* (2nd ed., pp. 487–499). Springer. https://link.springer.com/referenceworkentry/10.1007/978-0-387-30424-3_36
- Cummins, J. (2016). Reflections on Cummins (1980), “The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue.” *TESOL Quarterly*, 50(4), 940–944. <http://www.jstor.org/stable/44984725>
- King, K. & Bigelow, M. (2018). The language policy of placement tests of newcomer English learners. *Educational Policy* 32(7), 936–968. <https://doi.org/10.1177/0895904816681527>.
- Kudo, M., & Swanson, H. L. (2014). Are there advantages for additive bilinguals in working memory tasks? *Learning and Individual Differences*, 35, 96–102. <https://doi.org/10.1016/j.lindif.2014.07.019>
- Lau v Nichols, 414 U.S. 563 (1974). <https://www2.ed.gov/about/offices/list/ocr/ell/lau.html>
- Loose, J. (n.d.). Teaching language & content to ELLs. *More than English*. <https://morethanenglish.edublogs.org/for-teachers/wida-english-language-development-standards/>.
- Lopez, A. & Garcia, G. (2020). Teachers’ perspectives on the use of summative English language proficiency assessments for instructional purposes. In M. Wolf (Ed). *Assessing English Language Proficiency in U.S. K–12 Schools*. (pp. 226–243). Taylor & Francis.

- Makarova, E., Gilde, J., & Birman, D. (2019). Teachers as risk and resource factors in minority students' school adjustment: An integrative review of qualitative research on acculturation. *Intercultural Education, 30*(5), 448–477. <https://doi.org/10.1080/14675986.2019.158621>
- McEachern, F.M. (2022). A history of bilingual education in the US: Examining the politics of language policymaking. *International Journal of Bilingual Education and Bilingualism, 25*(9), 3526–3530. <https://doi.org/10.1080/13670050.2022.2078653>
- Paradis, J. (2019). English second language acquisition from early childhood to adulthood: The role of age, first language, cognitive, and input factors. In *Proceedings of the Boston University Conference on Language Development, 43*, pp. 11–26).
- Park, M., O'Toole, A., & Katsiaficas, C. (2017). *Dual language learners: A demographic and policy profile of South Carolina*. Migration Policy Institute.
- Plyler v Doe, 457 U.S. 202 (1982). <https://supreme.justia.com/cases/federal/us/457/202/>
- Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Sage.
- Scarcella, R. (2003, April). *Academic English: A conceptual framework*. <https://escholarship.org/content/qt6pd082d4/qt6pd082d4.pdf>
- South Carolina Department of Education. (n.d.). *Writing component*. from <https://ed.sc.gov/tests/middle/sc-ready/writing-component/>
- United States v Texas, 321 F. 1043 (5th Cir. 1970). https://www.supremecourt.gov/opinions/22pdf/22-58_i425.pdf
- Varga, S., & Guignon, C. (2020, February 20). *Authenticity*. Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/entries/authenticity/>
- Waters, C. N. (2020). *Teachers' perceptions of the broad validity of a high stakes English language proficiency test* [Dissertation, Virginia Commonwealth University]. Bilingual, Multilingual, Multicultural, and Education Commons. <https://scholarscompass.vcu.edu/cgi/viewcontent.cgi?article=7262&context=etd>.
- WIDA. (2020). *WIDA English language development standards framework, 2020 edition: Kindergarten–grade 12*. Board of Regents of the University of Wisconsin System. <https://wida.wisc.edu/sites/default/files/resource/WIDA-ELD-Standards-Framework-2020.pdf>
- WIDA (2023, April 28). *South Carolina English language learner identification and placement guidance document*. South Carolina Department of Education. <https://wida.wisc.edu/sites/default/files/id-placement/SC-ID-Placement-Guidance.pdf>
- Wisconsin Center for Education Research. (2023). *Access for ELLs*. WIDA. <https://wida.wisc.edu/assess/access>

Appendix A

Table 1

Legal Developments in Support of ML Education

Legal Development	Court Ruling	Implications for MLs
<i>Bilingual Education Act</i> (1968)	The segregation of non-native English speaking students in public general education classrooms is unlawful.	This court ruling was the first federal acknowledgment of MLs in public education. The ruling in favor of ML education also established federal grants to support ML education programs.
<i>United States v Texas</i> (1970)	An annual review of school districts by measuring the growth of MLs' language development is required in all public educational institutions.	This ruling established progress monitoring standards to track English language development and resulted in the widespread use of annual standardized testing to measure language proficiency growth.
<i>Lau v Nichols</i> (1974)	Identical education does not present equal learning opportunities for all students.	Public education must take affirmative action to remove educational barriers for all students and schools are to be held accountable for ensuring all students receive appropriate education.
<i>Casteñada v Pickard</i> (1981)	All instructional strategies must be based on legitimate educational theories and supported by existing research.	Schools must implement instructional programs, resources, and personnel necessary to provide appropriate education to all students. Educators must also be appropriately qualified and knowledgeable.
<i>Plyler v Doe</i> (1982)	Public schools cannot deny education to undocumented immigrants.	This ruling ensures education for all students despite background, culture, or home language. The passing of this ruling also resulted in steadily increasing ML student populations in public schools.

(Table adapted from Ariza & Coady, 2018)

Appendix B

WIDA Language Domains and Expectations

Table B1

Language Domains

Domains	Time Allotted	Expected Tasks	Answer Format	Scoring Type
Reading	60 minutes	-View a picture -Read a passage -Read the question -Click the correct answer	Selected response answers	Automatic Scoring
Writing	90 minutes	-View a picture -Read a short passage -Read the question -Write an appropriate response	Constructed response -Grades 1-3 paper response -Grades 4-5 typed response	Triangulated scoring by independent, trained test delivery partners
Listening	65 minutes	-View a picture -Listen to a voice talk -Listen to the question -Click or drag the answer	Selected response answers	Automatic Scoring
Speaking	50 minutes	-Listen to a brief exchange -Listen to the question -Record answer using a microphone	Constructed oral response recorded on a computer	Triangulated scoring by independent trained test delivery partners

(Table adapted from Coulter, 2017; Wisconsin Center for Education Research, 2023)

Table B2*Language domain expectations relating to English proficiency levels*

Proficiency	Listening	Speaking	Reading	Writing
Entering	Understand oral messages that include visuals and gestures and contain everyday words in English	Communicate orally using gestures and language that may contain a few words	Understand written texts that include visuals and may contain a few words or phrases	Communicate in writing using visual and symbols that may contain few words
Beginning	Understand oral language related to specific familiar topics and participate in class discussions	Communicate ideas and information orally using short sentences and everyday words and phrases	Understand written language related to specific familiar topics and participate in class discussions	Communicate in writing using language related to familiar topics
Developing	Understand oral language related to specific common topics and participate in class discussions	Communicate ideas and details orally using several connected sentences and participate in short conversations and discussions	Understand written language related to common topics and participate in class discussions	Communicate in writing using language related to common topics
Expanding	Understand oral language related to specific topics and participate in class discussions	Communicate orally using language related to specific topics in school and participate in class discussions	Understand written language related to specific topics	Communicate in writing using language related to specific topics
Bridging	Understand oral language and participate in all academic classes to recall examples and expand on ideas	Communicate orally and participate in all academic classes to discuss, summarize, present, and defend ideas	Understand written language from all academic classes to compare and summarize ideas and information	Communicate in writing using language from all academic classes
Reaching	Understand oral language and participate in all academic classes to synthesize information and recognize differing viewpoints	Communicate orally and participate in all academic classes to react, respond, persuade, and clarify multiple viewpoints	Understand written language from all academic classes to evaluate and synthesize ideas and information	Communicate in writing using language from all academic classes to produce clearly organized writing pieces of varying functions

Table adapted from WIDA (Wisconsin Center for Education Research, 2023)