Using the Woodcock-Johnson IV in Emergent Bilingual Evaluation of Cognition and Oral Language: Perceptions and Perspectives from the Field

Elizabeth A. Hatala, *Department of Education, West Texas A&M University*
Michelle Simmons, *Department of Education, West Texas A&M University*
Mikyung Shin, *Department of Special Education, Illinois State University*

Abstract

Selection confidence in assessment batteries varies among special education evaluators of culturally/linguistically diverse (CLD) and emergent bilingual (EB) students. In this cross-sectional, mixed-methods study, U.S. public school evaluators (n = 257) in multiple western states were surveyed about preferences and practices when using the Woodcock-Johnson IV Tests of Cognitive Abilities and the Woodcock-Johnson IV Tests of Oral Language with their CLD/EB caseloads. Results highlight perceptions of the battery's efficacy in circumstances of emergent bilingualism and multicultural considerations, concerns regarding the cultural bias of the instrument, and common practices when using this instrument for evaluation that are responsive to cultural and linguistic diversity.

Keywords: Bilingual evaluation, special education assessment, cultural linguistic diversity, text mining, educational diagnostician

The emergence of the profession of bilingual special education evaluator springs from a heightened need in many states for more culturally and linguistically sensitive evaluation of emergent bilingual (EB) or culturally and linguistically diverse (CLD) students referred for special education services (DeLeon & Gonzalez,1991). The National Center for Education Statistics (NCES) reported the percentage of public-school students who are identified as EB/CLD students in the United States increased from Fall 2011 to 2021 to 10.6 percent or 5.3 million enrolled students (2024). Among these national statistics, the state of Texas reported the highest percentage of public-school students who were EB/CLD students in comparison to all other states, with 20.2 percent of public enrolled students identified (National Center for Education Statistics, 2024).

Although there is a growing body of literature regarding the overrepresentation of EB/CLD students in special education programs (Garcia & Ortiz, 2004; Linan-Thompson, 2010; Huang et al., 2011; Sanatullova-Allison & Robison-Young, 2016), there is little written about the small-but-growing group of bilingual special education evaluators and how they manage the challenges of EB/CLD students in special education programs. The Texas Legal Framework defines special education evaluators as a "licensed specialist in school psychology (LSSP)/school psychologist, an educational diagnostician, or other appropriately certified or licensed practitioner with experience and training in the area of the disability" (Texas Administrative Code, 2024). For the purposes of this study, the term "special education evaluator" refers specifically to the role of licensed specialist in school psychology (LSSP)/school psychologist and educational diagnostician.

Studies of bilingual special education evaluators and student support teams serving a caseload of EB/CLD students report that many evaluators feel low levels of self-efficacy in making accurate identifications and intervention recommendations (Becker & Deris, 2019; Kritikos, 2003). To identify EB/CLD students effectively and reliably for special education services, bilingual special education evaluators must be familiar with second-language acquisition patterns that affect how a student performs both in the classroom and on standardized psycho-educational evaluation instruments (Cole et al., 2019). Knowledge of multilingual assessments and familiarity with multilingual learning characteristics are identified as recommended training for bilingual special education evaluators working with EB/CLD students. Another recommended area of focus in bilingual special education evaluator training is knowledge of testing to determine language proficiency in English and the student's native language (Cole et al., 2019). Bilingual special education evaluators' knowledge, training, and confidence in selecting assessments for evaluating EB/CLD students is critical.

Previous Studies on Culturally Sensitive Evaluation

Samuel Ortiz introduced a framework of best practices for nondiscriminatory assess-

Author Note

This research was funded by the Richard and Mary West Traylor Research Grant in Learning Disabilities, in association with the West Texas A&M University Center for Learning Disabilities. Our survey instrument was subjected to expert panel review by Joseph W. Madaus, Ph.D., Director of the University of Connecticut Collaborative on Postsecondary Education and Disability; W. Joel Schneider, Ph.D., of Temple University Department of Psychological Studies in Education; and Mary Bush Thomas of Texas Education Agency's Region 16 Education Service Center Multilingual Learner Support Services. We gratefully acknowledge these organizations and individuals for the financial and academic generosity of their contribution to this research effort.

Correspondence regarding this article should be directed to Elizabeth A. Hatala, 2800 Socorro Ct., College Station, TX, 77845. Email: eahatala1@buffs.wtamu.edu

ment of CLD learners (2002). Among points that focus on formal testing is the need for the assessment professional to be able to "test for intervention." In other words, the testing results should not only help decide eligibility for special education services but also provide enough instructionally informative data to guide intervention approaches for the child's progress, regardless of the eligibility outcome. Ortiz (2002) identified a framework of fundamental evaluator assets for nondiscriminatory evaluation (2002). As a follow-up study, Chen and Lindo (2018) reviewed 25 articles regarding best practices for evaluating CLD learners and outlined various aspects of cultural competence based on this foundational framework. Both studies stress that special education evaluators must have adequate knowledge of multilingualism and multiculturalism to select the most appropriate evaluation instrument, addressing both reliability and validity concerns (Ortiz, 2002; Chen & Lindo, 2018). Among the culturally competent behaviors highlighted was the use of culture as a lens through which behavior and test results should be interpreted (Ortiz, 2002; Chen & Lindo, 2018). Another notable element was adherence to IDEA-prescribed nondiscriminatory evaluation practices such as using multiple sources of data, a multidisciplinary team approach to decision-making, evaluating in the student's native language, and selecting a valid, nondiscriminatory assessment (Chen & Lindo, 2018). Using a reliable measure, one which is consistent across evaluators or across repeated measurements is crucial.

The Woodcock-Johnson (4th Ed.) Tests of Cognitive Abilities (WJ-IV COG), and the Woodcock-Johnson (4th Ed.) Tests of Oral Language (WJ-IV OL) are among several assessment batteries that have been created and normed to include features tailored to evaluating Spanish-speaking students. The WJ-IV OL has 9 tests in English and 3 tests in Spanish for easy comparison between performance in the two languages. It is marketed as ideal for determining English/Spanish proficiency and establishing levels of an EB/CLD student's Cognitive Academic Language Proficiency, or CALP (WPS, 2023a). However, a systematic review of comprehensive language assessments for monolingual children aged 4-12 found that the WJ-IV OL had not been the subject of any published peer-reviewed studies that were definitive measures of psychometric properties (Denman et al., 2017). The WJ-IV COG has a Spanish-language counterpart in La Batería IV Woodcock-Muñoz: Pruebas de habilidades cognitivas (La Batería-IV WM COG), also marketed as providing "an excellent comparison of skills when used in conjunction with the WJ-IV" (WPS, 2023b). The WJ-IV COG has been hailed as psychometrically sound by several studies (Bulut et al., 2021; Reynolds & Niileksela, 2015) while others point out weaknesses in the instrument (Canivez & Madle, 2017; Schneider, 2016). Published studies can be found that examine bias and invariance in the WJ-IV COG in school-age children across racial and ethnic groups (Gentry et al., 2021; Hajovsky & Chesnut, 2022; Izumi et al., 2019; Leahy Devine, 2020; Woods et al., 2021), with mixed results. Although the cross-cultural psychometrics of the WJ-IV COG and WJ-IV OL have been explored, the professional perspectives of multicultural assessors or practicing special education evaluators who use these tools regularly with students of cultural and linguistic diversity is a voice unheard.

Purpose of Study

This study surveyed the views held by bilingual special education evaluators primarily located in the state of Texas regarding the WJ-IV COG and the WJ-IV OL as reliable and valid measures for evaluating EB/CLD students referred for special education evaluation. Since many school districts may not have the benefit of employing a bilingual special education evaluator, especially rural schools in the state of Texas (Simmons, et al., 2020), this study aims to provide insights into the test instrument perspectives for monolingual, English-speaking evaluators during assessment instrument selection, administration, and results interpretation. The research study addressed the following research questions:

1. What are bilingual special education evaluators' perceptions of the efficacy of the WJ-IV COG, La Batería-IV WM COG, and the WJ-IV OL as reliable measures of language and cognition in the evaluation of emergent bilingual students?
2. What are bilingual special education evaluators' perceptions of cultural bias in the WJ-IV COG, La Batería-IV WM COG, and WJ-IV OL?

Method

Participants

As detailed in **Table 1**, respondents who passed the screening criteria of the online survey (n = 257) were certified educational diagnosticians or licensed specialists in school psychology with at least one academic year

PHILOSOPHY

The Texas Educational Diagnosticians' Association promotes professionalism, competence, excellence, and continuing education within the organization to provide individuals with disabilities quality educational services in the state of Texas.

of experience in the field who self-reported regularly evaluating EB/CLD students. Most (n = 194, 75.5%) respondents reported having an M.A., M.Ed., or equivalent degree. There were 72 (28%) evaluation personnel who had no teaching experience, suggesting they were most likely school psychologists (Zweback & Mortenson, 2002). Of respondents who reported teaching experience, 21% had five years of experience or less, while the remaining reported six years of experience or more. Of the 185 respondents with teaching experience, 59 (31.9%) reported no special education teaching experience, and 118 (63.8%) reported no bilingual teaching experience. Survey respondents were employed by public school districts primarily in the western United States. The greatest number were employed in Texas (n = 163, 63.4%), with 24 (9.3%) respondents in Utah, 24 (9.3%) in New Mexico, and 18 (7%) employed in California. Most respondents were Caucasian (n = 205, 79.8%), 17 (6.6%) were African American, and 15 (5.8%) were biracial or multiracial. More than half of the respondents (n = 158, 61.5%) across all races reported no Hispanic/Latinx ethnicity. Most respondents were between the ages of 41-50 years old (n = 101, 39.5%), and 64 (25%) reported ages 31-40.

Instrumentation

Instrumentation (see **Figure 1**) was based on existing surveys of cultural and linguistic diversity responsiveness considerations in peripheral fields such as speech-language pathology, clinical diagnosis of autism, teacher preparation, and school psychology (Harris et al., 2014; Kritikos, 2003; Mahalingappa, 2023; Vega et al., 2016) as the framework for design. Select survey questions were adapted to reflect the focal shift to special education evaluator self-efficacy using the WJ-IV COG and WJ-IV OL in CLD evaluation circumstances. The survey contained 57 questions and aimed to take 15-20 minutes to complete. At the end of the survey, participants had the option to provide contact information for follow-up questions or for future research participation. The Qualtrics online survey platform was used to create and later distribute the survey electronically. Survey content validity was verified through expert panel evaluation by professionals not associated with the study (Presser & Blair, 1994). The expert panel pretested the survey before it was approved for use. Their feedback was used to drop, add, and modify questions.

Table 1 Respondent Employment Location, Professional Experience, and Education

Characteristic	Respondents (n)	%
Teaching Experience		
None	72	28
5 yrs. or fewer	54	21
6-10 yrs.	57	22
11-15 yrs.	41	16
16-20 yrs.	20	7.8
21 yrs. or more	13	5.1
Special Ed. Teaching Experience		
None	59	31.9
5 yrs. or fewer	54	29.2
6-10 yrs.	37	20
11-15 yrs.	23	12.4
16-20 yrs.	12	6.5
Bilingual Ed. Teaching Experience		
None	118	63.8
5 yrs. or fewer	20	10.8
6-10 yrs.	24	13
11-15 yrs.	11	5.9
16-20 yrs.	9	4.9
21 yrs. or more	3	1.6
State of Employment		
Texas	163	63.4
New Mexico	24	9.3
Utah	24	9.3
California	18	7
Other States	28	10.9
Level of Education		
Master's Degree (MA, M.Ed.)	194	75.5
Specialist Degree (SSP, Ed.S.)	40	15.6
Doctorate (Ph.D, Ed.D, Psy.D.)	23	8.9

Note. Totals (n = 257) for each domain may not add up to 257 participants (100% response rate) due to non-reporting of data or survey skip-logic based on previous responses.

Figure 1 Principal Points of the Instrumentation and Data Collection Process

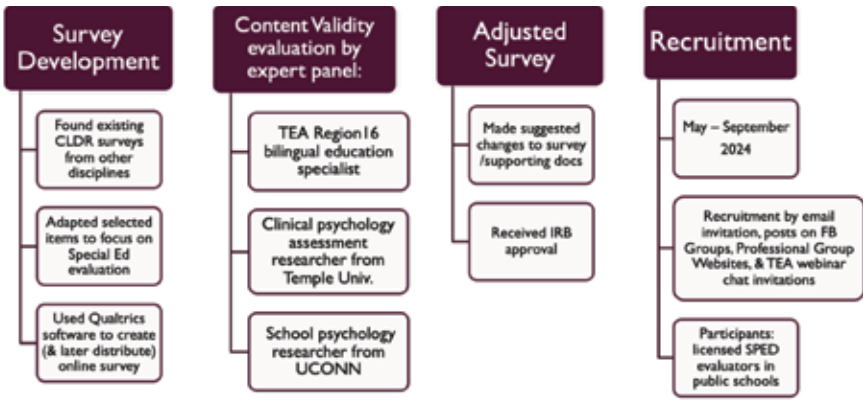


Table 2 User Efficacy Ratings of WJ-IV COG, Batería-IV WM COG, WJ-IV OL with EB/CLD Students

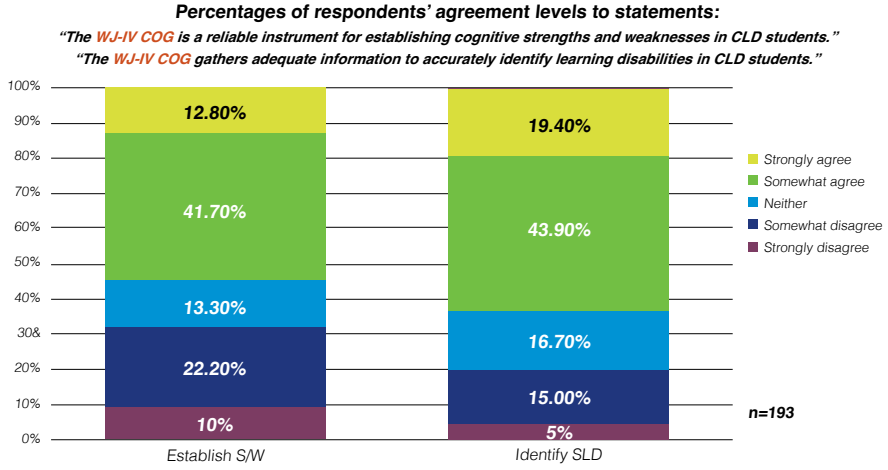
Perspective Statement	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
The WJ-IV COG					
-is reliable for finding cognitive strengths/weaknesses in EB/CLDs.	18 (10%)	40 (22.2%)	24 (13.3%)	75 (41.7%)	23 (12.8%)
-gathers enough data to identify learning disabilities.	9 (5%)	27 (15%)	30 (16.7%)	79 (43.9%)	35 (19.4%)
La Batería-IV WM COG					
-is reliable for finding cognitive strengths/weaknesses in "newcomer" Spanish-speaking students.	8 (6.2%)	14 (10.9%)	17 (13.2%)	63 (48.8%)	27 (20.9%)
-is reliable for finding cognitive strengths/weaknesses in Spanish-speaking students with some English schooling.	4 (3.1%)	15 (11.6%)	21 (16.3%)	63 (48.8%)	26 (20.2%)
-gathers enough data to identify learning disabilities.	3 (2.3%)	14 (10.9%)	30 (23.3%)	58 (45%)	24 (18.6%)
The WJ-IV OL					
-is reliable for finding language dominance.	3(1.9%)	24(14.9%)	20(12.4%)	77(47.8%)	37(23%)
-accurately measures CALP (cognitive academic language proficiency).	4(2.5%)	13(8.3%)	27(17.2%)	80(51%)	33(21%)
-gathers reliable supplementary data to identify some learning disabilities.	1 (0.6%)	13 (8.1%)	32(19.9%)	86 53.4%)	29 (18%)

Note. Number of respondents varies based on use of the test in question.

Data Collection and Analysis

Participants were recruited over five months via email invitation forwarded by school district special education directors, direct email invitation, posts on social media professional groups, professional organization websites, and chat invitations posted during a TEA webinar event. The researchers employed mixed methods for data analysis (see **Figure 1**). To understand respondents' perceptions and perspectives using the WJ-IV, the researchers used descriptive analysis and text-mining approaches. First, descriptive statistics were generated using jamovi version 2.3.28 (https://www.jamovi.org/) open-source software (The jamovi project, 2023). Microsoft Excel software was used to create visualizations of 5-point Likert scale responses (see **Figures S1-S5**). As a secondary analysis, we conducted word network analysis and analyzed the words associated with open-ended responses from survey respondents. For the text mining process, *TextAnalysisR* (https://textanalysisr.org) web application (Shin, 2024) was used. After importing the Excel file, researchers selected relevant columns that included targeted open-ended questions (e.g.,"What are the reasons for not using WJ-IV COG?") to preprocess text data. The following procedures were completed during this preprocessing (constructing a corpus, tokenizing text, converting to lowercase, removing stop words, and setting minimum length as token as two characters). We constructed a document-feature matrix and analyzed word correlation networks between two pairs of co-occurring words within each response.

Figure S1



Results

Efficacy of the WJ-IV COG, La Batería-IV WM COG, and WJ-IV OL

Descriptive statistics results show bilingual special education evaluators' perceptions of the efficacy of the WJ-IV COG and WJ-IV OL. Respondents shared perceived efficacy and reliability by rating perspective statements regarding the instrument (see **Table 2**). Of the respondents who had used the WJ-IV COG (n = 193), 98 (54%) agreed that it was a reliable measure for establishing cognitive strengths and weaknesses in CLD students, and 114 (63.3%) reported that it gathered adequate information to accurately identify learning disabilities in CLD students (see Figure S1). Similar questions were used regarding La Batería-IV WM COG, used by 132 (55.7%) respondents who tested Spanish-speaking students (See **Figure S2**). Of those respondents, 82 (63.6%) agreed that La Batería-IV WM COG gathered adequate information to

determine learning disabilities in CLD students. Respondents rated La Batería-IV WM COG equally reliable to establish cognitive strengths and weaknesses both in Spanish-speakers with no previous schooling in English (n = 90, 69.7%) and in emergent bilingual students with previous English exposure (n = 89, 69%).

There were 161 respondents who reported experience using the WJ-IV OL (see **Figure S3**). As detailed in **Table 3**, over 70% responded positively on the reliability of the instrument ***for*** determining language dominance (n = 114,71%) and gathering the supplementary data for identifying learning disabilities (n = 115, 71.4%). The majority of respondents (n = 110, 72%) also responded with positive views of the accuracy of the WJ-IV OL for measuring cognitive academic language proficiency (CALP).

Cultural Bias of the WJ-IV COG, La Batería-IV WM COG and WJ-IV OL

The second research question addressed the level of perceived cultural bias of the WJ-IV. When asked about the cultural bias of the WJ-IV COG (see **Figure S4**), 65 (35%) respondents regarded the scripted instructions as culturally unbiased and only 37 (20%) perceived test items to be culturally unbiased (see **Table 3**). When similar questions addressed perceptions of the WJ-IV OL, results indicated that 67 (42.2%) of the 152 participating respondents perceived no cultural bias in the test items (see **Figure S5**).

Text mining results showed evaluator concerns with items in the WJ-IV COG, WJ-IV OL, and Batería-IV related to cultural conflicts. With the minimum word co-occurrence of 3 within each response, we selected word pairs with correlation coefficients greater than 0.2. As shown in **Figure 2**, the word "culturally" maintained the highest degree centrality, connecting the strongest associations with other words such as "outdated," "language," and "heavy." Other clustered words such as "culture," "difficult," and "understanding" also reflected concerns related to cultural issues addressed in the administered items. Other concerns included the fact that tests are based on "English" and "vocabulary." One respondent pointed out, "When I used to use the WJ, it always had very low scores for my Speech and ELL [English Language Learner] students. I learned early on that if students had weaker language and vocabulary skills, it resulted in low scores in many of their (sub)tests."

Additionally, some respondents shared that they did not use WJ-IV COG, WJ-IV OL, and Batería-IV, with EB/CLD students. Based

Figure S2

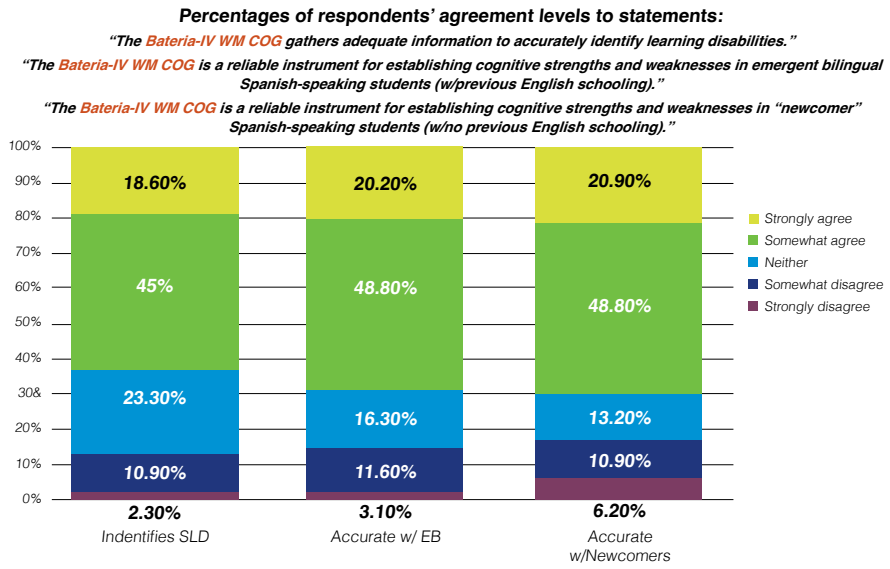


Figure S3

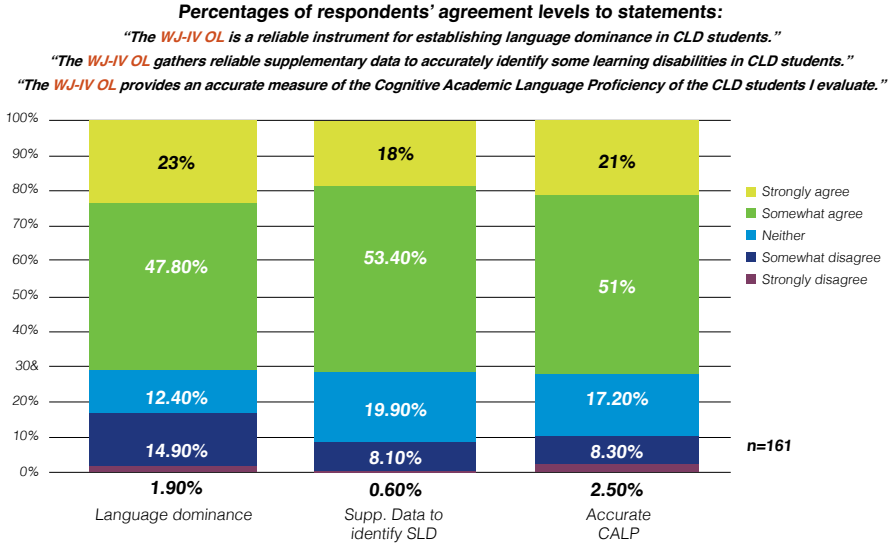


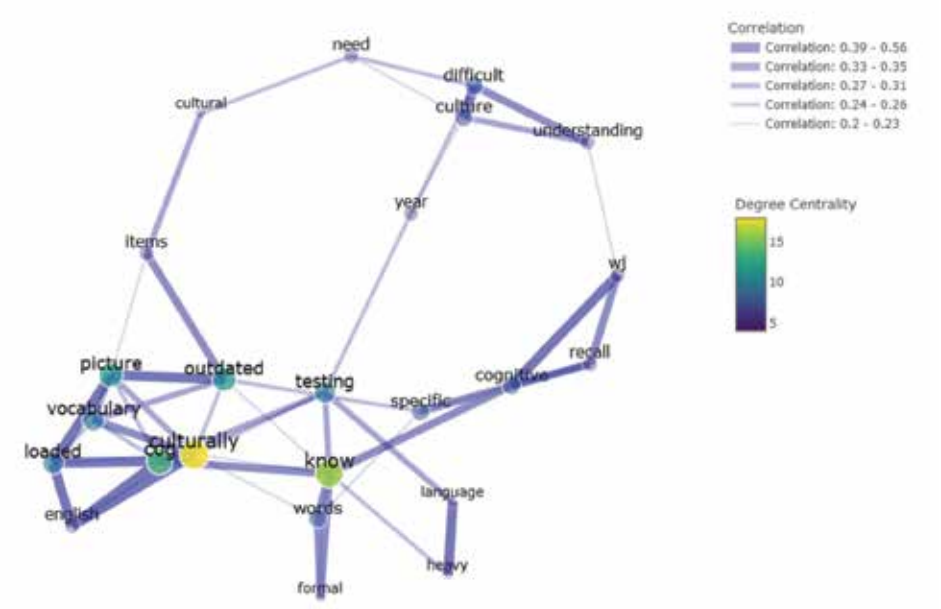
Table 3 Respondent Ratings of Cultural Bias in WJ-IV COG, WJ-IV OL with EB/CLD Students

Perspective Statement	Strongly Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Strongly Agree
The WJ-IV COG					
-scripted instructions are culturally neutral enough.	23(12.4%)	65(35.1%)	32(17.3%)	54(29.2%)	11(5.9%)
-test items are free from cultural bias.	28(15.1%)	83(44.9%)	37(20%)	32(17.3%)	5(2.7%)
The WJ-IV OL					
-test items are free from cultural bias.	14(8.8%)	38(23.9)	40(25.2%)	58(36.5%)	9(5.7%)

Note. Number of respondents varies based on use of the test in question.

on the minimum word co-occurrence results of 3 within each response and at least 0.2 between pairwise word correlation, respondents' shared responses of not using the assessments were intertwined. As shown in Figure 3, the words "evaluation" and "cognitive" showed the two highest degrees centrality, connecting other related words around them. One consistent concern was related to "language" or "loaded" aspects of the assessments. The centered word of "evaluation" points out "SLP" and "concerns." These associations addressed some respondents' opinions, such as "I do not use any OL. Language concerns are usually addressed by either an SLP or a bilingual school psych on staff in my district." One other respondent shared that the reason for not using WJ-IV COG was that "The district I work for requires a Wechsler Nonverbal to be done." The word network demonstrates this connection of "nonverbal" evaluation instead of the language-based assessment for EB/CLD students.

Figure 2 Word Correlation Network on the WJ-IV COG, WJ-IV OL, and Batería-IV Examples of Cultural Conflicts Note. Minimum co-occurrence numbers = 3; minimum correlation between pairwise words = 0.2.



Discussion and Conclusions

Analysis of bilingual special education evaluators' responses to the survey reveals several important findings that can help in conducting future EB/CLD special education evaluations. Although previous studies indicated many evaluators of multilingual/multicultural students feel low levels of self-efficacy in making accurate identifications and intervention recommendations for EB/CLD students (Becker & Deris, 2019; Kritikos, 2003), most respondents were confident in the WJ-IV COG as a reliable measure. WJ-IV COG users reported more concerns about cultural bias than other reliability factors.

According to current special education evalu-

Figure 3 Word Correlation Network on Not Using WJ-IV COG, WJ-IV OL, and Batería-IV With CLD Students *Note. Minimum co-occurrence numbers = 3; minimum correlation between pairwise words = 0.2.*

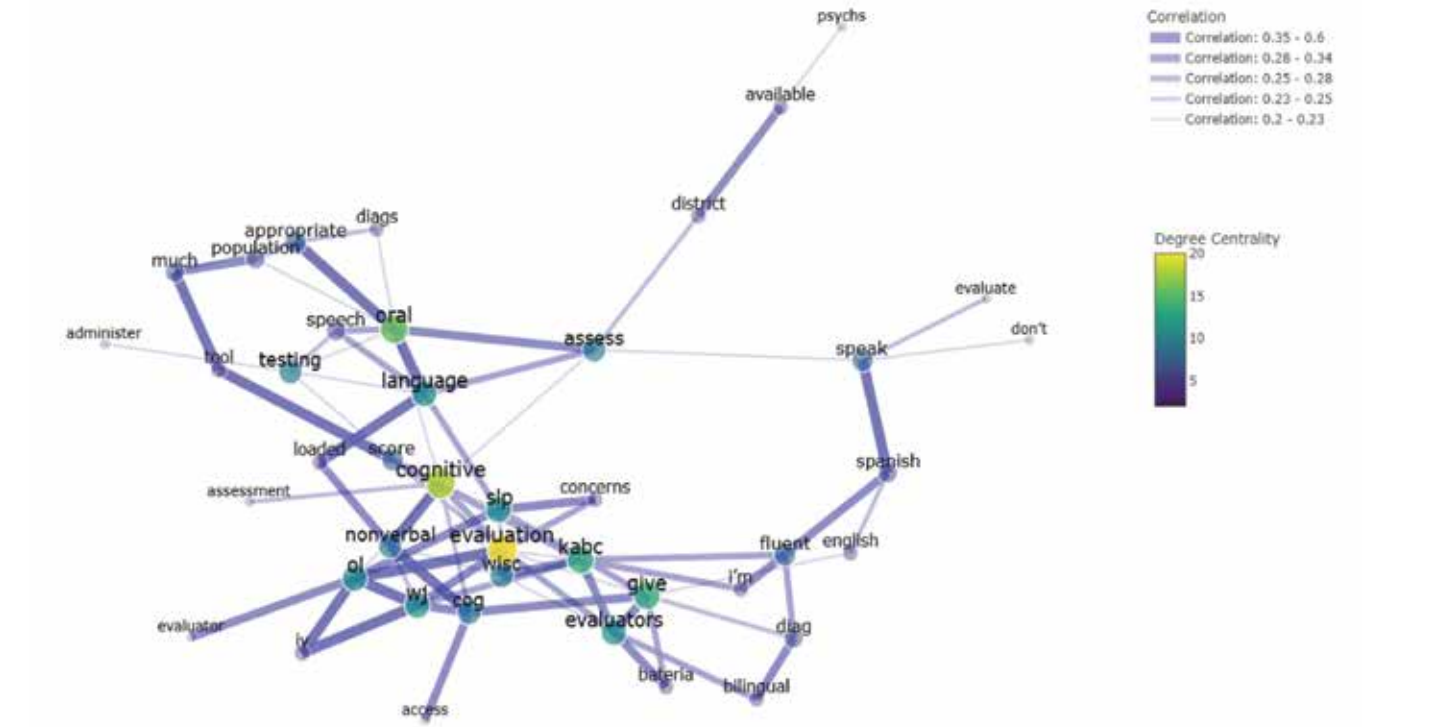


Figure S4

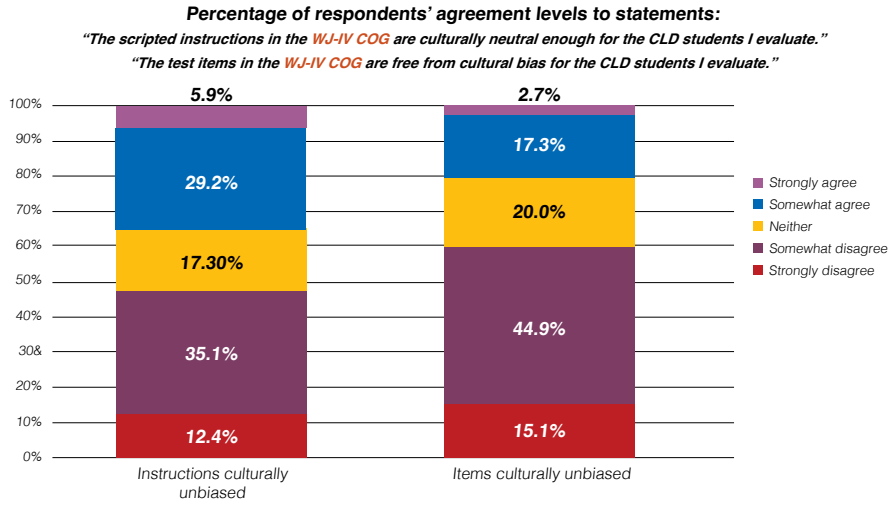


Figure S5

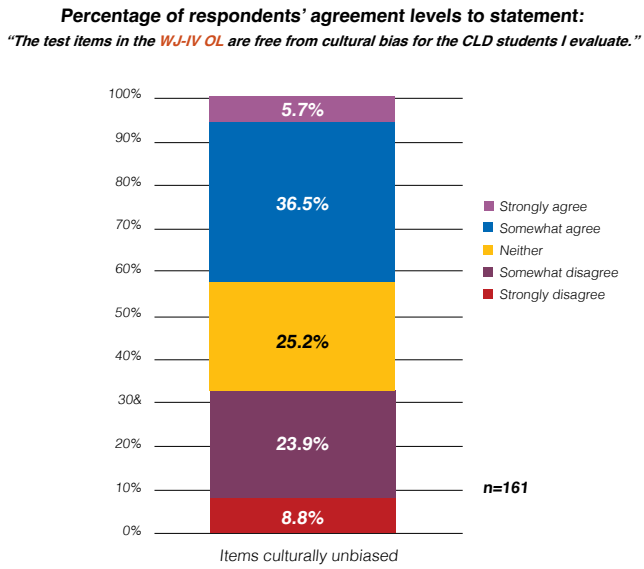


Table 4 Supplemental Assessment Batteries Recommended

WJ-IV Tests of Cognitive Abilities	WJ-IV Tests of Oral Language
WISC-V	WMLS
KABC-II	WMLS-III
WISC-V Spanish	PVAT
	WIAT-4
	BVAT

ators, the WJ-IV OL is a reliable measure for determining language dominance and gathering supplementary data for identifying learning disabilities, and an accurate measure of cognitive academic language proficiency. Overall, respondents were more confident with the use of the WJ-IV OL. Respondents, however, also referenced several other preferred assessments in addition to the WJ-IV COG and WJ-IV OL when evaluating EB/CLD students (see **Table 4**).

The current study indicated methodological benefits that education researchers could replicate in their mixed-method studies. To validate the perspectives of practitioners, we employed both quantitative (using Likert-scale questions) and qualitative (through written responses) approaches, thereby deepening our understanding of different voices across multiple sources and related questions. Text mining techniques could support the normalization of large text data and efficiently detect associations among frequently co-occurring words. Future researchers should further validate the process and explore how the text-mining approach can enhance the analysis of open-ended and qualitative data in the field.

This research distills the professional experiences of 257 licensed or certified special education evaluators who regularly assess EB/CLD students in U.S. public schools. The anticipated impact is that these results provide support to evaluators who are less experienced with this special population, especially those in rural or small school districts where the likelihood of finding a bilingual evaluator is greatly reduced. Continued research in this area can positively shape the special education evaluation and service delivery experienced by EB/CLD students with learning disabilities.

References

Becker, G. I, & Deris, A. R. (2019). Identification of Hispanic English language learners in special education. *Education Research International*, vol. 2019, Article ID 2967943, 9 pp. https://doi.org/10.1155/2019/2967943

Bulut, O., Cormier, D. C., Aquilina, A. M., & Bulut, H. C. (2021). Age and sex invariance of the Woodcock-Johnson IV tests of cognitive abilities: Evidence from psychometric network modeling. *Journal of Intelligence, 9*(3), 1–16. https://doi-org.databases.wtamu.edu/10.3390/jintelligence9030035

Canivez, G. L., & Madle, R. A. (2017). Woodcock-Johnson® IV. *The Twentieth Mental Measurements Yearbook.*

Chen, C. R., Lindo, E. J. (2018). Culturally responsive practices in special education evaluations: A review of literature. *DiaLog: Journal of the Texas Educational Diagnosticians Association. 47*(2), 9-13.

Cole, C. V., Blackwell III, W., & Duran, J. B. (2019). A university program to prepare educational diagnosticians to engage in culturally and linguistically responsive evaluation practices for English learners. *Journal of universality of global education issues.* https://ugei-ojs-shsu.tdl.org/ugei/article/view/44

De Leon, J., & Gonzales, E. (1991). An ex-

amination of bilingual special education and related training. *Teacher Education and Special Education, 14*(1), 5–10. https://doi.org/10.1177/088840649101400102

Denman, D., Speyer, R., Munro, N., Pearce, W.M., Chen, Y-W., and Cordier, R. (2017). Psychometric Properties of Language Assessments for Children Aged 4–12 Years: A Systematic Review. *Frontiers in Psychology, 8*,1515. https://doi.org/10.3389/fpsyg.2017.01515

Garcia, S. B., & Ortiz, A. A. (2004). Preventing inappropriate referrals of language minority students to special education. National Clearinghouse for English Language Acquisition.

Gentry, M., Desmet, O. A., Karami, S., Lee, H., Green, C., Cress, A., Chowkase, A., & Gray, A. (2021). Gifted education's legacy of high stakes ability testing: Using measures for identification that perpetuate inequity. *Roeper Review: A Journal on Gifted Education, 43*(4), 242–255. https://doi-org.databases.wtamu.edu/10.1080/02783193.2021.1967545

Hajovsky, D. B., & Chesnut, S. R. (2022). Examination of differential effects of cognitive abilities on reading and mathematics achievement across race and ethnicity: Evidence with the WJ IV. *Journal of School Psychology, 93*, 1–27. https://doi-org.databases.wtamu.edu/10.1016/j.jsp.2022.05.001

Harris, B., Barton, E. A., Albert, C. (2014). Evaluating autism diagnostic and screening tools for cultural and linguistic responsiveness. *Journal of Autism and Developmental Disorders*, [s.l.], 44(6), 1275–1287. https://doi.org/10.1007/s10803-013-1991-8

Huang, J., Clarke, K., Milczarski, E., & Raby, C. (2011). The assessment of English language learners with learning disabilities: Issues, concerns, and implications. *Education, 131*(4) pp. 732+. Accessed 20 Sept. 2023, from Gale Academic OneFile, link.gale.com/apps/doc/A260137595/AONE?u=txshracd2629&sid=bookmark-AONE&xid=e173b01d

Izumi, J. T., Burns, M. K., & Frisby, C. L. (2019). Differences in specific learning disability identification with the Woodcock-Johnson IV. *School Psychology, 34*(6), 603–611. https://doi-org.databases.wtamu.edu/10.1037/spq0000336

The jamovi project. (2023). jamovi (Version 2.3.28) [Computer Software]. https://www.jamovi.org

Kritikos, E. P. (2003). Speech-language pathologists' beliefs about language assessment of bilingual/bicultural individuals. American *Journal of Speech-Language Pathology, 12*(1), 73-91. https://doi.org/10.1044/1058-0360(2003/054)

Leahy Devine, M. (2020). Does analysis of linguistic characteristics of the WJ IV directions predict degree of linguistic demand? [ProQuest Information & Learning]. In Dissertation Abstracts International Section A: Humanities and Social Sciences (Vol. 81, Issue 8–A).

Linan-Thompson, S. (2010). Response to instruction, English language learners and disproportionate representation: The role of assessment. Psicothema, 22(4), 970-974. https://www.proquest.com/scholarly-journals/response-instruction-english-language-learners/docview/2778377364/se-2

Mahalingappa, L. (2023). Building teacher candidates' self-efficacy beliefs about the education of multilingual learners: A linguistically responsive approach, *Teaching and Teacher Education, 131*. https://doi.org/10.1016/j.tate.2023.104190.

National Center for Education Statistics. (2024). English learners in public schools. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. https://nces.ed.gov/programs/coe/indiocator/cgf

Ortiz, S. O. (2002). Best Practices in Nondiscriminatory Assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology IV* (pp. 1321–1336). National Association of School Psychologists.

Presser, S. & Blair, J. (1994). Survey pretesting: Do different methods produce different results? Sociological Methodology, 24, 73-104.

Reynolds, M. R., & Niileksela, C. R. (2015). Review of Woodcock-Johnson IV Tests of Cognitive Abilities. *Journal of Psychoeducational Assessment, 3*3(4), 381–390. https://doi-org.databases.wtamu.edu/10.1177/0734282915571408

Sanatullova-Allison, E., & Robison-Young, V. A. (2016). Overrepresentation: An Overview of the Issues Surrounding the Identification of English Language Learners with Learning Disabilities. *International Journal of Special Education, 31*(2), 145–151. https://eric.ed.gov/?id=EJ1111073

Schneider, W. J. (2016). Strengths and weaknesses of the Woodcock-Johnson IV Tests of Cognitive Abilities: Best practice from a scientist-practitioner perspective. In D. P. Flanagan & V. C. Alfonso (Eds.), WJ IV clinical use and interpretation: Scientist-practitioner perspectives. (pp. 191–210). Elsevier Academic Press. https://doi-org.databases.wtamu.edu/10.1016/B978-0-12-802076-0.00007-4

Shin, M. (2024). TextAnalysisR: Text mining workflow tools [Web application]. https://www.textanalysisr.org

Simmons, M., Shin, M., Sharp, L. (2020). Special education eligibility identification rates in Texas: A comparative analysis of rural and urban school districts. *Texas Education Review, 9*(1), http://dx.doi.org/10.26153/tsw/11422

Texas Administrative Code, 19 Tex. Admin. Code § 89.1040(b)(1) (2024). https://texreg.sos.state.tx.us/public/readtac$ext.TacPage?app=9&ch=89&rl=1040&ti=19&pt=2&sc=AA&rl=1040

Vega, D., Lasser, J., & Afifi, A.F.M (2016). School psychologists and the assessment of culturally and linguistically diverse students. *Contemporary School Psychology, 20*(3), 218-229. https://doi.org/10.1007/s40688-015-0075-5.

Woods, I. L., Jr., Niileksela, C., & Floyd, R. G. (2021). Do Cattell-Horn-Carroll cognitive abilities predict reading achievement similarly for Black children as for other racial/ethnic groups? *Contemporary School Psychology, 25*(2), 183–199. https://doi-org.databases.wtamu.edu/10.1007/s40688-020-00338-1

WPS. (2023a). (WJ IV) Woodcock-Johnson IV. https://www.wpspublish.com/wj-iv-woodcock-johnson-iv.html

WPS. (2023b). Batería IV Woodcock-Muñoz. https://www.wpspublish.com/bateria-iv-woodcock-munoz.html

Zweback, S., & Mortenson, B. P. (2002). State Certification/Licensure Standards for Educational Diagnosticians: A National Review. *Education, 123*(2), 370–379.

Supplemental Items

Participants

Other states were represented to a lesser degree, such as Louisiana's seven (2.7%) respondents, the five (1.9%) from Arizona, four (1.6%) from Oklahoma, three (1.2%) from Colorado, and one (0.38%) lone respondent from each of the following states: NV, OR, KS, NJ, CT, FL, MA, WI, NE, NY.

Additional races reported were five (1.9%) respondents who claimed Native American or Alaskan Native heritage, and two (0.8%) were Asian or Pacific Islander."Other race" was selected by 13 (5.1%) respondents.