



WJ V Nonverbal: Development, Validation, and Interpretive Guidance

Evelyn S. Johnson, EdD

Hyeonjoo Oh, PhD

Huan Liu, PhD

Nicole Roviso, MEd

The Woodcock-Johnson® V Nonverbal (WJ V™ Nonverbal) was developed to reduce the linguistic demands of cognitive assessment while preserving the constructs measured by selected WJ V cognitive tests. Administration uses visual demonstrations, standardized gestures, and reduced verbal mediation while retaining the original items, scoring procedures, and test structure. This bulletin describes the development of the WJ V Nonverbal and presents evidence supporting interpretation of the WJ V Nonverbal Cognitive Ability Index (NVCAI).

In a counterbalanced study of 475 school-age examinees, NVCAI scores were closely similar across verbal and nonverbal administration for general education students, English learners, and students with language impairments, with no significant Version × Group interaction. Additional analyses showed strong concurrent validity with the KABC-II NU Nonverbal Index ($r = .765$) and high NVCAI reliability across the age span ($r = .96-.97$). Confirmatory factor analysis also provided support for the proposed CHC-based structure. Together, the findings support use of the WJ V Nonverbal as a language-minimized component of comprehensive psychoeducational evaluation.

Copyright © 2026 by Riverside Assessments, LLC. All rights reserved. No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or by any information storage or retrieval system without the prior written permission of Riverside Assessments, LLC. Permission is hereby granted to individuals having permissible access to these materials to photocopy entire pages from this publication for personal, non-commercial use and not for resale or any commercial activity. Requests for information on other matters regarding duplication of this work should be sent via email to permissions@riversideinsights.com or addressed via mail to Riverside Insights, Attention: Permissions, One Pierce Place, Suite 900W, Itasca, Illinois 60143.

Riverside Insights, the Riverside Insights logo, Riverside Score, and Woodcock-Johnson are registered trademarks of Riverside Assessments, LLC. WJ V, the WJ V logo, WJ V Nonverbal, and WJ V NV are trademarks of Riverside Assessments, LLC.

IBM and SPSS are registered trademarks of International Business Machines Corporation.

iPad is a registered trademark of Apple, Inc.

KABC and Kaufman Assessment Battery for Children are trademarks of NCS Pearson, Inc.

For technical information, please call 800.343.9540 or visit our website at https://www.riversideinsights.com/woodcock_johnson_v

Reference Citation

- To cite this document, use:

Johnson, E. S., Oh, H., Liu, H., & Roviso, N. (2026). *WJ V Nonverbal: Development, validation, and interpretive guidance*. Riverside Assessments, LLC.

WJ V Nonverbal: Development, Validation, and Interpretive Guidance

Introduction

The WJ V Nonverbal (WJ V NV) was developed to minimize the linguistic demands of cognitive assessment while preserving the cognitive constructs measured by the original *Woodcock-Johnson V Tests of Cognitive Abilities* (WJ V COG). The composite score of the WJ V Nonverbal is the WJ V Nonverbal Cognitive Ability Index (NVCAI) and is composed of seven subtests drawn from the WJ V COG. The WJ V Nonverbal is intended to support equitable assessment practices for individuals whose performance may be affected by language-based barriers rather than limitations in cognitive functioning.

The need for such an instrument reflects a longstanding challenge in psychoeducational assessment. Many traditional measures of cognitive ability rely heavily on verbal comprehension, linguistic knowledge, and spoken instructions. While such assessments yield valuable information, the linguistic demands embedded in cognitive test batteries represent a potential source of construct-irrelevant variance, where test performance reflects language proficiency rather than the cognitive construct being measured (Cormier et al., 2011). This concern is especially salient in contemporary educational and clinical settings, where practitioners routinely evaluate English learners (EL), students with developmental language disorders, students with hearing impairments, and individuals from culturally and linguistically diverse backgrounds (McCallum, 2017).

The WJ V Nonverbal was designed to address this need directly. Its administration procedures employ visual demonstration, gesture-based communication, and reduced verbal mediation, while preserving the original task demands, scoring procedures, and interpretive framework of the underlying assessments. This is not a claim to language-free or culture-free measurement. Rather, the WJ V Nonverbal reduces unnecessary linguistic demands so that an examinee's performance more accurately reflects the cognitive abilities under examination. The result is a complementary set of tools that supports more equitable and interpretable evaluation, particularly for examinees whose language proficiency or linguistic background might otherwise introduce systematic bias into test performance. This bulletin describes the conceptual foundation, development process, administration procedures, and research supporting the WJ V Nonverbal and the NVCAI, with the goal of providing practitioners a clear understanding of how it was developed, how the term “nonverbal” is operationalized within the assessment process, and how scores should be interpreted within a comprehensive evaluation framework.

Structure and Constructs of the WJ V Nonverbal

A key conceptual distinction underlying the WJ V Nonverbal is the difference between truly nonverbal assessment and reduced-language assessment. Truly nonverbal approaches attempt to eliminate verbal stimuli and verbal responses entirely. Reduced-language approaches retain some verbal components but systematically lower the linguistic complexity of instructions, prompts, and response demands. In practice, most contemporary nonverbal cognitive batteries occupy a position along this continuum. The WJ V Nonverbal is best understood as a language-minimized measure—one that preserves the psychometric integrity of its underlying constructs while reducing the linguistic demands that can obscure cognitive functioning in examinees for whom language is not a reliable indicator of ability.

The WJ V Nonverbal includes seven subtests. Six of the subtests form three broad abilities within the Cattell-Horn-Carroll (CHC) theory of cognitive abilities: Fluid Reasoning (Gf), Visual Processing (Gv), and Cognitive Processing Speed (Gs). The seventh subtest measures Visual Working Memory. These subtests were selected because they measure cognitive functioning relevant to psychoeducational evaluation and can be administered through highly visual, demonstration-based procedures with minimal reliance on spoken language. Subtest selection was guided by two criteria: theoretical alignment with CHC broad ability domains and suitability for administration through visual, demonstration-based procedures with minimal verbal mediation. The following descriptions of the three broad abilities, their associated subtests, and the Visual Working Memory test as they appear within the WJ V Nonverbal have been adapted from the WJ V Technical Manual (LaForte et al., 2025).

Fluid Reasoning (Gf)

The Fluid Reasoning (Gf) cluster contains the Matrices and Analysis-Synthesis tests. Fluid reasoning is the use of deliberate and controlled procedures to solve novel, “on the spot” problems that cannot be solved by relying on previously learned habits, schemas, and scripts. Fluid reasoning is often considered a hallmark indicator of general intelligence.

Visual Processing (Gv)

The Visual Processing (Gv) cluster is comprised of the Block Rotation and Spatial Relations tests. Visual processing can be defined as the ability to make use of simulated mental imagery to solve problems—perceiving, discriminating, manipulating, and recalling nonlinguistic images in the “mind’s eye.” Gv is one of the most psychometrically studied CHC broad ability domains and has received increased neurocognitive interest related to preparing individuals in the STEM fields.

Cognitive Processing Speed (Gs)

The Cognitive Processing Speed (Gs) cluster is comprised of the Number-Pattern and Letter-Pattern Matching tests. Cognitive processing speed is the ability to control attention to perform both simple and complex repetitive cognitive tasks automatically, quickly, and fluently. Gs also may be described as attentional fluency or attentional speediness.

Visual Working Memory

The Visual Working Memory test is a mixed measure of Gwm and Gv. It assesses an individual’s ability to recall a visual-spatial pattern of stimuli in the context of processing (i.e., after performing a simple visual distractor task).

Development of Nonverbal Administration Procedures

Development of the WJ V Nonverbal administration procedures involved two related components: modifying how directions were communicated to examinees and establishing standardized guidance for examiners to ensure consistent implementation. Each is described below.

Modification of Directions

The WJ V Nonverbal was developed to reduce the linguistic demands of administration while preserving the underlying cognitive constructs being measured. Test items, scoring criteria, basal and ceiling rules, and item sequencing remained unchanged from the standardized forms of the assessments. Modifications were limited to the way task expectations were communicated to the examinee. To support these modifications, test developers systematically reviewed the original task demands, examiner instructions, and response requirements for each subtest and translated the verbal directions into functionally equivalent nonverbal forms. Across subtests, the primary modification strategy was the development of wordless video demonstrations that illustrated task directions and expectations using sample and demonstration items.

The video-based directions were refined through repeated internal review and piloting. In the initial development phase, test developers reviewed each demonstration to determine whether it communicated the required task expectations clearly. The revised directions were then piloted with 20 students, including English learners and students with language-based disabilities. Student feedback and observed areas of confusion were used to identify unclear or ambiguous elements, and the procedures were revised accordingly.

Development of Standardized Examiner Guidance

Because examiner behavior can directly affect the validity of nonverbal assessment, the WJ V Nonverbal provides explicit guidance for standardized administration. Examiners are expected to use only the video-based directions, standardized gestures, and allowable prompts specified for each subtest. This consistency is especially important because the procedures rely on standardized visual communication rather than spoken directions. Even well-intentioned behaviors, such as redirecting an examinee's attention, emphasizing stimulus features, repeating a demonstration beyond what is specified, or otherwise supplementing the directions, may inadvertently influence performance and complicate score interpretation. For this reason, the administration procedures distinguish allowable prompts and gestures from examiner behaviors that fall outside the standardized procedures and that should be avoided. Detailed guidance for each subtest is provided directly within the *Riverside Score*® platform.

Research Study 1: Comparing WJ V Verbal and Nonverbal Versions

A primary concern in the development of any nonverbal adaptation of an existing cognitive assessment is whether modified administration procedures alter the construct being measured or systematically affect performance relative to the original standardized form. To address this concern, a research study was conducted to evaluate whether the WJ V Nonverbal produces scores that are comparable to those obtained under standard WJ V verbal administration conditions and whether any observed differences vary across examinees with different linguistic profiles. The study was designed to address the following research question: Do the nonverbal and verbal administration procedures of the selected WJ V tests operate similarly for general education students, students with language impairments, and English learners?

Hypotheses regarding the expected pattern of results were informed by the theoretical framework for evaluating the validity of testing accommodations (Johnson, 2000; Sireci & Faulkner-Bond, 2015; Thurlow & Kopriva, 2015). Within this framework, an accommodation is considered valid when it produces score changes for students for whom the standard administration introduces construct-irrelevant barriers (i.e., students who require accommodation), while leaving scores unchanged for students who do not require accommodation. Applied to the present study, this framework generates distinct predictions for each participant group. For general education students, nonverbal and verbal administration procedures were expected to yield comparable scores, as these students are unlikely to be disadvantaged by the linguistic demands of standard administration. For students with language impairments (LI) and English learners, by contrast, nonverbal administration was hypothesized to result in scores that were equal to or higher than verbal administration scores, reflecting the removal of linguistic barriers that may otherwise suppress performance on the standard form. A finding consistent with this pattern would provide evidence both that the nonverbal procedures preserve the underlying cognitive constructs and that they function as a meaningful accommodation for examinees whose language abilities may not accurately represent their cognitive functioning.

Participants

Examiner Recruitment and Training

Data were collected by a cadre of 32 professional examiners possessing specialized expertise in educational and psychological assessment. The examiner cohort comprised school psychologists, educational diagnosticians, licensed school counselors, and special education professionals, all of whom had experience in the administration of standardized psychometric instruments.

Examiners were recruited based on their assessment experience and their ability to recruit EL and LI students. Recruitment focused on professionals with experience in research, including prior participation in Riverside Insights norming studies such as the WJ V standardization. Additional outreach targeted school psychologists and other assessment professionals in regions with relatively high enrollments of EL or LI students to help meet the study's sampling goals.

Prior to data collection, all examiners completed standardized training to support consistent administration of the WJ V assessments using verbal and nonverbal directions. The asynchronous online modules covered the *Riverside Score* platform, standardized testing procedures, test content, study-specific data collection requirements, and the differences between the verbal and nonverbal administration conditions. The training was intended to promote consistency across examiners and reduce errors, procedural drift, and variation in how the two conditions were implemented.

Before collecting data, examiners completed a full practice case in the *Riverside Score* platform. They administered all six study subtests in sequence under both conditions, beginning with the verbal protocols and then completing the corresponding nonverbal protocols. The practice case served as a fidelity check of each examiner's ability to navigate the platform, capture responses, score accurately, and implement each administration condition correctly. Examiners were cleared to begin participant recruitment after they successfully completed the full practice case.

Participants and Sampling Design

The study used a balanced, *a priori* sampling design targeting a total sample of 510 school-aged examinees across three cohorts: a general education control group, an EL group, and an LI group, with 170 examinees targeted per cohort. Participants included children and adolescents from Kindergarten through Grade 12. Grade-level recruitment targets were organized into four developmental bands: Grades K–2, Grades 3–5, Grades 6–8, and Grades 9–12. Recruitment targets for gender and race/ethnicity were based on federally validated reference benchmarks. Targets for the general education and LI cohorts were based on U.S. Census Bureau school-age population estimates. Because the demographic profile of English learners differs from that of the general school-age population, targets for the EL cohort were based on national public school enrollment data from the National Center for Education Statistics, specifically the Fall 2015 EL demographic distributions reported in *Status and Trends in the Education of Racial and Ethnic Groups* (de Brey et al., 2019).

Sample Distribution

The final sample closely aligned with the projected sampling framework and included 475 examinees: 171 in the general education control group, 155 in the LI group, and 149 in the EL group. Table 1 summarizes the sample distribution by developmental grade band. Broader demographic characteristics included sex, race/ethnicity, study group, and Free or Reduced-Price Lunch (FRPL) eligibility. FRPL eligibility was used as a proxy for socioeconomic status (SES).

Table 1.
*Demographic Characteristics
of the Final WJ V Nonverbal
Analysis Sample by Study
Group*

Demographic Variable	Study Group			
	General Education (<i>n</i> = 171)	Language Impaired (<i>n</i> = 155)	English Learner (<i>n</i> = 149)	Total (<i>N</i> = 475)
Grade Band				
K–2	41 (24.0)	41 (26.5)	44 (29.5)	126 (26.5)
3–5	40 (23.4)	43 (27.7)	37 (24.8)	120 (25.3)
6–8	46 (26.9)	40 (25.8)	30 (20.1)	116 (24.4)
9–12	44 (25.7)	31 (20.0)	38 (25.5)	113 (23.8)
Sex				
Female	87 (50.9)	76 (49.0)	86 (57.7)	249 (52.4)
Male	84 (49.1)	79 (51.0)	63 (42.3)	226 (47.6)
Race/Ethnicity				
White	90 (52.6)	80 (51.6)	0 (0.0)	170 (35.8)
Black/African American	25 (14.6)	26 (16.8)	12 (8.1)	63 (13.3)
American Indian or Alaska Native	1 (0.6)	12 (7.7)	1 (0.7)	14 (2.9)
Asian	10 (5.8)	1 (0.6)	16 (10.7)	27 (5.7)
Native Hawaiian/Other Pacific Islander	1 (0.6)	0 (0.0)	0 (0.0)	1 (0.2)
Two or More Races	12 (7.0)	9 (5.8)	6 (4.0)	27 (5.7)
Hispanic or Latinx	32 (18.7)	27 (17.4)	114 (76.5)	173 (36.4)
Free or Reduced-Price Lunch				
Yes	79 (46.2)	60 (38.7)	95 (63.8)	234 (49.3)
No	75 (43.9)	50 (32.3)	35 (23.5)	160 (33.7)
Unsure/Not Reported	17 (9.9)	45 (29.0)	19 (12.8)	81 (17.1)

Note. Values are presented as *n* (%). Percentages are calculated within each study group.

Participants in the general education control group had no active IEP/504 plan or special education services, although gifted-only students were eligible. Participants in the EL group were required to be actively receiving English learner services at the time of testing; students receiving only indirect language support or accommodations were excluded. The LI group included students actively receiving special education services for speech-language impairment, deafness, hearing impairment, or a documented language-based learning disability. Across all groups, students with severe disabilities, including autism spectrum disorder Level 3, intellectual disability, emotional disturbance, traumatic brain injury, or developmental delay, were excluded.

Design and Procedure

The study employed a counterbalanced, within-subjects design in which each participant completed both the verbal and nonverbal versions of the WJ V across two separate administrations. Data were collected between January and June 2026 across all four major U.S. Census geographic regions (Midwest: *n* = 37, South: *n* = 214, Northeast: *n* = 121, West: *n* = 103).

All assessments were administered using a dual-device configuration within the *Riverside Score* platform. The examiner used a laptop to launch the testing session and view procedural directions, while the student interacted directly with a linked iPad® that automatically captured and scored responses in real time. Test sessions were conducted individually in a quiet, private setting free from external distractions. A minimum interval of 2 days was required between administrations to reduce the potential influence of fatigue and immediate practice effects.

Administration order was counterbalanced across participants within each group so that approximately half of the students in each group received the verbal form first and half received the nonverbal form first, ensuring that any order effects were distributed evenly across conditions. For verbal administrations, examiners followed the standard directions specified in the WJ V COG administration procedures without modification. For nonverbal administrations, examiners followed the video-based directions and standardized gestures as specified.

Data Analysis

The analysis evaluated whether the WJ V Nonverbal tests produced scores comparable to those obtained under standard verbal administration and whether any score differences between versions varied across the three examinee study groups. The dependent variable was the NVCAI standard score, obtained under the verbal and nonverbal administration conditions. It was computed by averaging the *W* ability scores of six tests, Matrices (MATRCZ), Spatial Relations (SPAREL), Number-Pattern Matching (NUMPAT), Analysis-Synthesis (ANLSYN), Block Rotation (BLKROT), and Visual Working Memory (VWKMEM), and converting the resulting mean *W* score to a standard score using the age-based norm table. Although the published WJ V Nonverbal comprises seven tests, only six tests were administered. Because Letter-Pattern Matching (LETPAT) and NUMPAT share highly similar nonverbal instructions, data were collected for NUMPAT only to reduce examinee testing burden.

The NVCAI was analyzed using a three-way mixed analysis of variance (ANOVA) with Version (verbal vs. nonverbal) as the within-subjects factor and Group (general education, EL, and LI) and Order (verbal first vs. nonverbal first) as the between-subjects factors. The two effects of primary interest were the Version main effect, which evaluated overall score differences between the verbal and nonverbal versions, and the Version $\times$ Group interaction, which tested whether such differences were consistent across the three examinee groups. In addition, the Version $\times$ Order interaction was examined as an indicator of practice effects (i.e., whether scores increased as a function of administration order), and the Version $\times$ Group $\times$ Order interaction was examined to see whether any practice effect varied systematically across groups. Cohen's *d* is reported as the effect size for paired version differences, and the correlation between the verbal and nonverbal versions is provided as a measure of alternate-version consistency.

Prior to analysis, records were removed for invalid administration, including failure to complete both versions and too short an interval between the two testing sessions. Analyses were based on the examinees who completed all six tests under both versions.

Results

Table 2 presents the means (*M*), standard deviations (*SD*), and correlations (*r*) for the verbal and nonverbal NVCAI standard scores, by group and for the full sample. Mean scores under the verbal and nonverbal versions were similar within every group, with mean differences of less than 2 standard-score points. Scores on the verbal version were slightly higher in every group. The two versions were strongly correlated across all groups ($r = .73-.84$), indicating high alternate-version consistency. As expected, the three groups differed in overall performance level, with the general education group scoring highest and the LI group scoring lowest. These descriptive patterns were evaluated inferentially with a three-way mixed ANOVA (Version $\times$ Group $\times$ Order) on the NVCAI (see Table 3).

Table 2.
*Descriptive Statistics for
the NVCAI Composite
Under Verbal and Nonverbal
Administration*

Group	<i>N</i>	<i>M</i> Verbal	<i>M</i> Nonverbal	<i>M</i> Diff	<i>SD</i> Diff	<i>r</i>	Cohen's <i>d</i>
All	475	98.64	97.69	.95	8.54	.83	.11
General education	171	105.43	103.61	1.82	7.74	.83	.23
English learners	149	98.10	97.47	.62	9.50	.73	.07
Language impaired	155	91.66	91.36	.30	8.40	.84	.04

Table 3.
*Three-Way Mixed ANOVA
(Version $\times$ Group $\times$ Order) for
the NVCAI Composite*

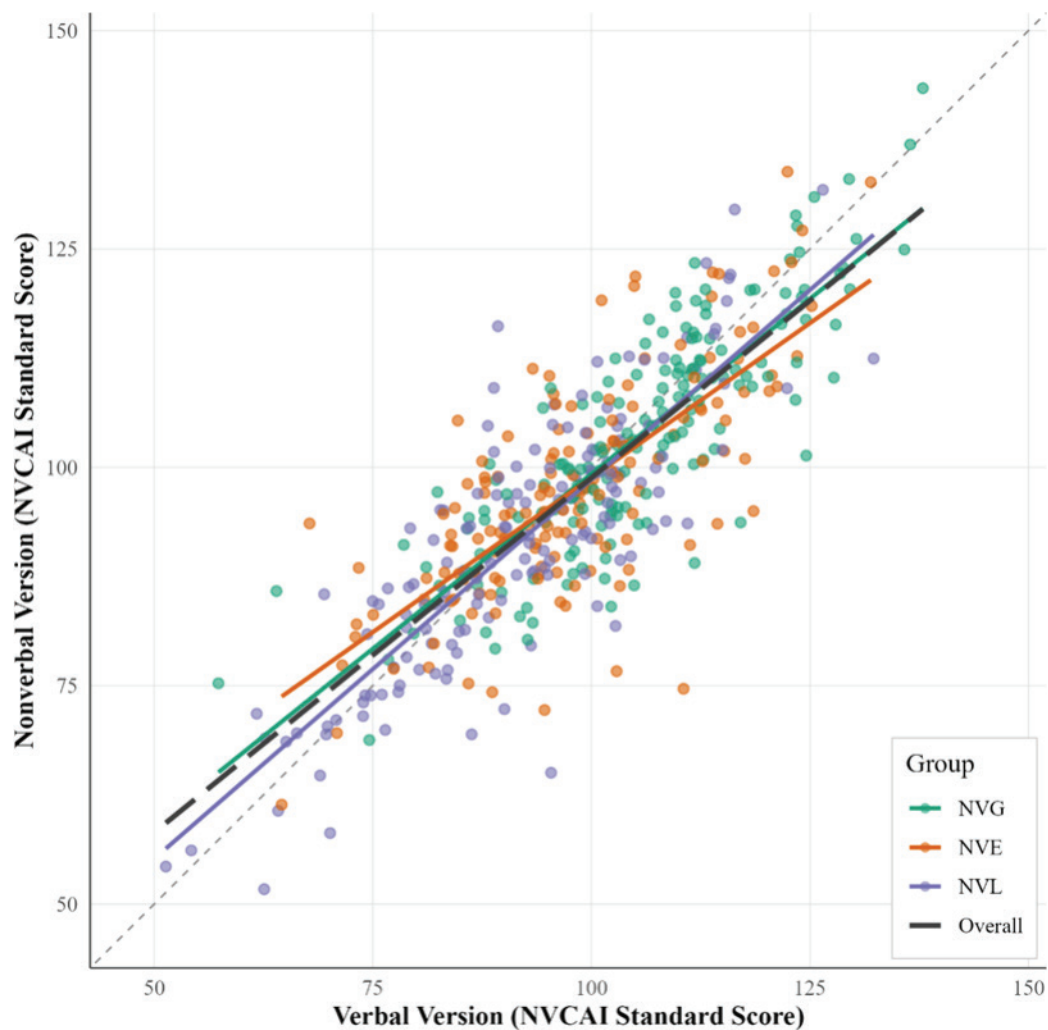
Effect	<i>df</i>	<i>F</i>	<i>p</i>
Group	2,469	41.71	< .001
Order	1,469	.94	.333
Group $\times$ Order	2,469	.60	.550
Version	1,469	11.84	< .001
Version $\times$ Group	2,469	.89	.410
Version $\times$ Order	1,469	194.33	< .001
Version $\times$ Group $\times$ Order	2,469	2.82	.061

Note. * indicates statistical significance at $p < .05$

Comparability of the Verbal and Nonverbal Versions

As indicated in Table 3, the ANOVA showed a small but significant main effect of Version, $F(1, 469) = 11.84$, $p < .001$, and no Version $\times$ Group interaction, $F(2, 469) = .89$, $p = .410$. These results indicate that performance was slightly higher under the verbal version, by about 1 standard-score point overall, and that this small difference held consistently across general education examinees, Els, and examinees with LI, with mean version differences of approximately 2 standard-score points or less in every group. Descriptively, the difference was largest among general education examinees and, as shown in the test and cluster-level analyses below, attributable primarily to a single test, Number-Pattern Matching (NUMPAT), within that group. As shown in Figure 1, scores under the two versions clustered closely around the line of identity across all three groups, further illustrating the overall close correspondence between versions.

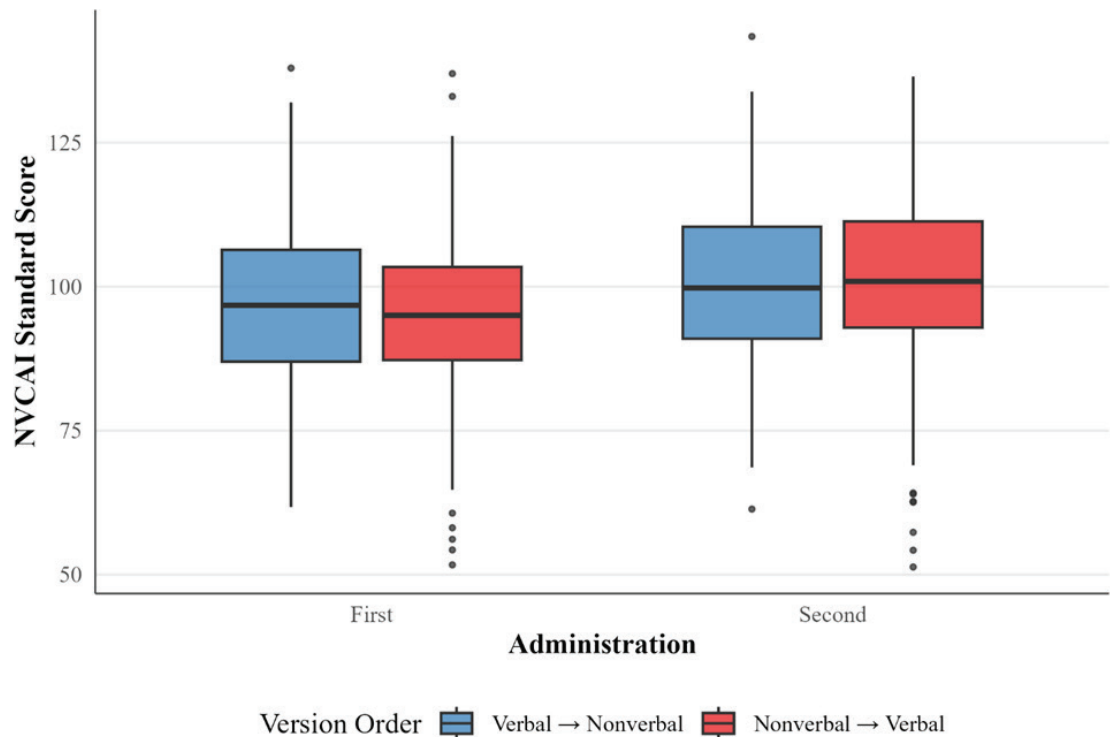
Figure 1.
Verbal Versus Nonverbal
NVCAI Standard Scores by
Examinee Group



Practice Effects

Although the two versions produced equivalent scores overall, Table 3 shows that a substantial practice effect was observed, reflected in a highly significant Version $\times$ Order interaction, $F(1, 469) = 194.33, p < .001$. This effect was attributable to testing occasion rather than version. Regardless of which version was administered first, scores on the second administration ($M = 100.45, SD = 14.88$) exceeded those on the first ($M = 95.87, SD = 14.06$) by approximately 5 standard-score points ($M \text{ Diff} = 4.58, SD \text{ Diff} = 7.27, d = .63$). Because order was counterbalanced, this practice gain was distributed evenly across the two versions and did not bias the comparison between them (Figure 2). An exploratory test of whether the practice effect differed across groups yielded a nonsignificant, three-way interaction, Version $\times$ Group $\times$ Order, $F(2, 469) = 2.82, p = .061$, suggesting that the practice effect was consistent across all three examinee groups.

Figure 2.
*NVCAI Standard Scores by
Administration and Version
Order*



Group and Order Effects

As expected, the three groups differed in overall ability level (see Table 3), $F(2, 469) = 41.71$, $p < .001$, with general education examinees scoring highest and the examinees with LI scoring lowest. The two order groups did not differ in overall ability, $F(1, 469) = .94$, $p = .333$, and this equivalence held within each group, $\text{Group} \times \text{Order}$, $F(2, 469) = .60$, $p = .550$, indicating that examinees were well balanced across the two administration orders with no evidence of selection bias.

In summary, the verbal and nonverbal versions produced closely similar scores across all three examinee groups. The Version main effect, although statistically significant, corresponded to a difference of about 1 standard-score point overall, and there was no Version $\times$ Group interaction. The practice effect of roughly 5 standard-score points was observed but, due to counterbalancing, did not affect the comparison between versions. Taken together, these results indicate that the nonverbal directions yield scores comparable to those obtained under the standard verbal administration across general education examinees, Els, and examinees with LI, with the small residual difference favoring the verbal version examined at the test level below.

Cluster and Test Scores

To examine whether the comparability of the two versions held across specific ability domains, the three-way mixed ANOVA (Version $\times$ Group $\times$ Order) used for the composite was applied to the two cluster scores (Gf and Gv) and the six tests. Descriptive statistics and alternate-version correlations are summarized in Table 4, and the ANOVA results are shown in Table 5. The pattern observed for the composite was largely reproduced: the focal Version $\times$ Group interaction was nonsignificant for seven of the eight outcomes (the exception being NUNPAT), the Version $\times$ Order interaction was significant for every outcome, and the three-way Version $\times$ Group $\times$ Order interaction was nonsignificant throughout. Neither cluster score showed an overall version effect (Gf, $p = .986$; Gv, $p = .418$). At the subtest level, the Version main effect was significant for two of the six subtests; the remaining four subtests showed no version effect. The two significant effects, however, followed different patterns across groups.

Table 4.

Descriptive Statistics and Alternate-Version Correlations for the WJ V Cluster and Test Scores

Outcome	N	M Verbal	M Nonverbal	M Diff	SD Diff	r	Cohen's d
Gf	475	99.68	99.97	-.29	13.80	.70	-.02
Gv	474	101.07	100.84	.23	10.54	.76	.02
MATRCZ	475	100.88	101.71	-.83	15.26	.64	-.05
ANLSYN	475	98.92	98.91	.01	15.57	.60	.00
SPAREL	475	101.53	101.64	-.12	11.17	.68	-.01
BLKROT	475	100.88	100.23	.65	12.62	.66	.05
NUMPAT	475	98.54	96.53	2.01	9.04	.77	.22
VWKMEM	475	100.89	99.56	1.33	14.18	.64	.09

Note. Gf = Fluid Reasoning, Gv = Visual Processing, MATRCZ = Matrices, ANLSYN = Analysis-Synthesis, SPAREL = Spatial Relations, BLKROT = Block Rotation, NUMPAT = Number-Pattern Matching, VWKMEM = Visual Working Memory

Table 5.

Three-Way Mixed ANOVA (Version × Group × Order) for the Cluster and Test Scores

Effect	V		V × G		V × O		V × G × O	
Outcome	F	p	F	p	F	p	F	p
Gf	.00	.986	.21	.807	101.61	< .001*	2.24	.108
Gv	.66	.418	.23	.795	43.08	< .001*	2.28	.103
MATRCZ	1.08	.300	.33	.716	24.49	< .001*	.24	.784
ANLSYN	.22	.637	.31	.735	75.80	< .001*	2.00	.137
SPAREL	.00	.978	.34	.711	28.02	< .001*	.93	.396
BLKROT	2.09	.149	.08	.921	23.01	< .001*	2.43	.089
NUMPAT	28.28	< .001*	4.74	.009*	62.70	< .001	.08	.920
VWKMEM	6.26	.013*	.37	.693	58.74	< .001*	.00	.996

Note. V = Version, G = Group, O = Order, Gf = Fluid Reasoning, Gv = Visual Processing, MATRCZ = Matrices, ANLSYN = Analysis-Synthesis, SPAREL = Spatial Relations, BLKROT = Block Rotation, NUMPAT = Number-Pattern Matching, VWKMEM = Visual Working Memory.

* $p < 0.05$

Number-Pattern Matching (NUMPAT), a processing speed test, showed a significant overall advantage over the verbal version of NUMPAT, $F(1,469) = 28.28$, $p < .001$ (see Table 5), amounting to about 2 standard-score points ($M \text{ Diff} = 2.01$, $d = .22$; see Table 4). The Version × Group interaction was nonsignificant, $F(2,469) = 4.74$, $p = .009$, indicating that the magnitude of the version effect differed across groups (see Table 5). Follow-up simple-effects tests shown in Table 6 indicated that this nonverbal advantage was statistically significant for the general education group and the EL group but not for the LI group. Although the NV NUMPAT version effect was statistically significant, its magnitude was small and of negligible practical consequence. The mean difference of approximately 2 standard-score points corresponds to a Cohen's d of .22 (see Table 4), essentially at the conventional threshold for a small effect (.20) and smaller than the test's standard error of measurement (SEM). A difference of this size would not produce meaningful differences in individual classification or interpretation.

Table 6.
Simple-Effects Comparisons
of WJ V Verbal Versus
Nonverbal by Group for
NUNPAT and VWKMEM

Test/Group	Estimate	SE	<i>t</i>	<i>p</i>
NUNPAT				
General education	3.71	.65	5.74	< .001*
English learners	1.00	.69	1.45	.148
Language impaired	1.50	.68	2.19	.029*
VWKMEM				
General education	.84	1.03	.82	.415
English learners	2.07	1.10	1.88	.060
Language impaired	1.73	1.09	1.60	.111

Note. Estimates are the mean difference (verbal – nonverbal) in standard-score points.

The verbal version of Visual Working Memory (VWKMEM) showed a smaller and more uniform pattern in the same direction: a slight advantage over the nonverbal version, $F(1,469) = 6.26$, $p = .013$ ($M\text{ Diff} = 1.33$, $d = .09$) (see Tables 4 and 5). This effect was uniform across groups (see Table 6). The verbal advantage was in the same direction in all three groups but did not reach significance within any single group ($p \geq .060$ for all groups), consistent with the lower power of within-group tests. Given the small magnitude of standardized difference, this effect is of limited practical significance.

In summary, at the composite level the two versions yielded essentially equivalent scores for English learners and for examinees with language impairments, the primary target populations for the nonverbal administration (mean NVCAI differences of .62 and .30 standard-score points, $d = .07$ and .04, respectively; see Table 2). In addition, the cluster scores supported comparability, whereas two individual tests showed small version-specific effects favoring the verbal version for both Number-Pattern Matching and Verbal Working Memory.

Research Study 2: Concurrent Validity

A second study was conducted to examine the psychometric validity of the WJ V Nonverbal by evaluating its relationship with an independently developed measure of nonverbal cognitive ability. Because the WJ V Nonverbal is theoretically grounded within the CHC framework, it is designed to isolate and measure specific CHC broad and narrow abilities while reducing verbal language demands. To demonstrate concurrent validity, scores from the WJ V Nonverbal clusters should correlate meaningfully with those obtained from other well-established nonverbal cognitive batteries. Specifically, this study evaluated the association between the WJ V Nonverbal and the Nonverbal Index of the *Kaufman Assessment Battery for Children®*, *Second Edition Normative Update* (KABC™-II NU; Kaufman & Kaufman, 2018) to determine the extent to which they assess overlapping nonverbal cognitive constructs independent of language proficiency.

Participants

The concurrent validity sample was nested within the broader WJ V Nonverbal study sample, with participants in this validation study drawn directly from the primary study sample. This nested approach ensured that the findings would reflect the clinical and linguistic populations targeted, while allowing for a direct comparison with an independent nonverbal cognitive measure. The concurrent validity sample consisted of ($N = 108$) school-aged examinees ages 6–18 who were enrolled in Kindergarten through Grade 12. This aggregate sample included the same three distinct participant groups established in the broader study: a general education population cohort ($n = 40$), a language impaired clinical cohort ($n = 40$), and an English learner cohort ($n = 28$).

The same inclusion and exclusion criteria established for Research Study 1 were applied uniformly to the concurrent validity sample. In addition to these criteria, participants were required to have valid, complete, and error-free scoring protocols for both the WJ V Nonverbal Cognitive Ability Index (NVCAI) and the KABC-II NU Nonverbal Index. Cases were excluded from the final analytic data set if they exhibited incomplete or invalid administrations, data-quality concerns, or failed to meet group-specific eligibility criteria. The demographic distribution of the final concurrent validity sample across the three target groups is presented in Table 7.

Table 7.
*Demographic Characteristics
of the Concurrent Validity
Sample by Study Group*

Demographic Variable	General Education (<i>n</i> = 40)	Language Impaired (<i>n</i> = 40)	English Learners (<i>n</i> = 28)	Total (<i>N</i> = 108)
Grade Band				
K–2	10 (25.0)	9 (22.5)	7 (25.0)	26 (24.1)
3–5	11 (27.5)	12 (30.0)	11 (39.3)	34 (31.5)
6–8	9 (22.5)	8 (20.0)	4 (14.3)	21 (19.4)
9–12	10 (25.0)	11 (27.5)	6 (21.4)	27 (25.0)
Sex				
Female	20 (50.0)	16 (40.0)	13 (46.4)	49 (45.4)
Male	20 (50.0)	24 (60.0)	15 (53.6)	59 (54.6)
Race/Ethnicity				
White	26 (65.0)	21 (52.5)	0 (0.0)	47 (43.5)
Black/African American	1 (2.5)	4 (10.0)	0 (0.0)	5 (4.6)
American Indian or Alaska Native	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Asian	1 (2.5)	0 (0.0)	0 (0.0)	1 (0.9)
Native Hawaiian/ Other Pacific Islander	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)
Two or More Races	5 (12.5)	5 (12.5)	0 (0.0)	10 (9.3)
Hispanic or Latinx	7 (17.5)	10 (25.0)	28 (100.0)	45 (41.7)
Free or Reduced-Price Lunch				
Yes	28 (70.0)	24 (60.0)	17 (60.7)	69 (63.9)
No	9 (22.5)	10 (25.0)	4 (14.3)	23 (21.3)
Unsure/Not Reported	3 (7.5)	6 (15.0)	7 (25.0)	16 (14.8)

Note. Values are presented as *n* (%). Percentages are calculated within each study group.

Measures

To evaluate the concurrent validity of the WJ V Nonverbal cluster, the KABC-II NU (Kaufman & Kaufman, 2018) was used as the external criterion measure. The KABC-II NU is a nationally standardized measure of cognitive ability designed for children ages 3–18. The KABC-II NU Nonverbal Index was selected as the criterion measure for this study because it was developed specifically to assess cognitive functioning in children for whom verbal responding may not be appropriate, including English learners, children with hearing impairments, and children with speech or language disorders. The internal consistency reliability coefficient of the KABC-II NU Nonverbal Index is .95 for students ages 7–18 (Kaufman et al., 2018).

Examiners

Although the broader WJ V Nonverbal study leveraged an extensive network of field examiners, data collection for the concurrent validity study was intentionally restricted to a specialized subset of four professional examiners. Because the valid administration of the external measure required

specialized training, these specific examiners were selected based on their verified proficiency with the KABC-II NU Nonverbal Index. Prior to active data collection, each examiner was sent a complete testing kit and manual with instructions to thoroughly review the standardized administration and scoring guidelines for the KABC-II NU Nonverbal Index. To monitor administration fidelity and protect against scoring errors, an internal quality-control protocol was maintained throughout the study. This involved conducting random, independent score audits on completed KABC-II NU protocols as they were submitted.

Design and Procedure

Data were collected between March and June 2026. While the parent study drew from a nationwide sample, participant recruitment and testing for this concurrent validity substudy were concentrated across sites in the Northeast (60.2%) and Southern (39.8%) regions.

Each participant completed both the WJ V, which included both the standardized WJ V and the WJ V Nonverbal tests counterbalanced according to the parent study's design. The administration order of the KABC-II NU Nonverbal Index was counterbalanced across these existing testing sessions to control for practice and order effects. Participants either received the KABC-II NU prior to their scheduled WJ V sessions or completed their WJ V administrations first before completing the KABC-II NU. This test battery was divided across multiple testing sessions to manage testing fatigue; administrations required a minimum of two sessions with at least 3 days of separation between them. In all cases, the entire testing sequence for any given examinee was completed within a 3-week window to control for developmental changes. For all administrations, examiners adhered strictly to the standardized procedures specified for each instrument. All testing protocols were scored using the standardized scoring criteria outlined in the respective test manuals, and the raw scores were subsequently converted to standard scores for statistical analysis.

Data Analysis

Concurrent validity was evaluated using Pearson product-moment correlations between WJ V Nonverbal Ability Index and KABC-II NU Nonverbal Index standard scores. Both instruments yielded standard scores with a normative mean of 100 and standard deviation of 15. All analyses were conducted in R (R Core Team, 2025).

Results

We recruited 126 examinees for the concurrent validity study, each of whom completed both the KABC-II NU Nonverbal and the WJ V Nonverbal tests, with administration order counterbalanced. On average, the two tests were administered approximately 2 weeks apart ($M = 15.24$ days, $SD = 18.83$) by trained examiners following standardized procedures. Consistent with Research Study 1, analyses were restricted to examinees who completed all six WJ V and WJ V NV tests. Because both sets of tests report standard scores only at or above 40, scores below this floor were not interpretable; therefore, examinees who scored below 40 on any test(s) were removed. These criteria yielded a final sample of 108 examinees. The two measures produced nearly identical mean score levels across the full sample (WJ V Nonverbal $M = 95.35$, $SD = 15.74$; KABC-II NU Nonverbal Index $M = 95.31$, $SD = 18.06$), a difference of only 0.04 standard-score points.

Across the full sample, a Pearson product-moment correlation indicated a strong, statistically significant association between the KABC-II NU Nonverbal Index and WJ V Nonverbal standard scores, $r(106) = .76$, $p < .001$, 95% CI (.67, .83), providing evidence of strong concurrent validity.

Reliability of the WJ V Nonverbal Cognitive Ability Index (NVCAI) Score

The NVCAI is a composite score formed from the WJ V Nonverbal tests. This analysis examined the reliability of the NVCAI by age across the full WJ V norming sample. NVCAI reliability was derived from the reliabilities and intercorrelations of its component tests, following the standard procedure for composite scores reported in the WJ V Technical Manual (LaForte et al., 2025). Reliability was estimated for the full sample (ages 6–80+), for school-age examinees (ages 6–19), and for adults (ages 20–80+) in Table 8 and for five age groups in Table 9.

Table 8.
*NVCAI Summary Statistics
and Reliability by Age Band*

Age Band	N	M	SD	<i>r</i> _{cc}	SEM (SS)
6–80+	5,159	100.75	10.64	.96	2.03
6–19	3,391	100.47	10.54	.96	2.05
20–80+	1,768	101.27	10.82	.97	1.98

Table 9.
*NVCAI Reliability by Age
Group*

Age Group	N	<i>r</i> _{cc}
6–9	938	.96
10–14	1,255	.96
15–19	1,198	.96
20–49	905	.97
50–80+	863	.97

Note. Median *r*_{cc} across age groups = .96

The standard error of measurement (*SEM*) was computed and reported in both the standard score metric and the W-score metric. The *SEM* indexes the expected variability of an observed score around the true score; smaller values indicate greater precision and yield narrower confidence intervals.

Results

Across the full norming sample (ages 6–80+, *N* = 5,159), the NVCAI was highly reliable (*r* = .96, *SEM* = 2.03 standard-score points). Reliability was comparable for the school-age band (ages 6–19, *r* = .96) and the adult band (ages 20–80+, *r* = .97), with *SEMs* near 2.0 standard-score points in both (see Table 8). Reliability was stable across the five age groups, with *r* ranging from .96 to .97 (see Table 9). The summary statistics and *SEM* by single year of age are summarized in Table 10. Taken together, the NVCAI demonstrates high reliability (*r* ≥ .96) across the entire age range, with small and stable *SEM* (approximately 2 standard-score points). These results indicate that the NVCAI provides precise and consistent measurement of nonverbal cognitive ability suitable for individual interpretation across the lifespan. Note that the reliabilities of the contributing tests and associated cluster scores are identical to those reported in the WJ V Technical Manual (LaForte et al., 2025) and are not reproduced in this bulletin.

Table 10.
*NVCAI Summary Statistics and
SEM by Single Year of Age*

Age	N	M	SD	SEM (W)	SEM (SS)
6	224	480.16	9.75	1.82	1.95
7	230	488.26	9.66	1.78	1.91
8	237	493.41	10.00	1.81	1.95
9	247	499.19	11.09	1.85	1.99
10	247	503.84	11.06	1.87	2.00
11	251	507.71	12.61	1.94	2.06
12	253	511.35	12.47	1.96	2.06
13	250	514.87	13.68	2.01	2.09
14	254	517.71	13.05	2.03	2.07
15	243	520.37	14.24	2.07	2.09
16	243	519.43	14.08	2.06	2.12
17	230	523.51	14.32	2.11	2.11
18	251	525.23	13.57	2.13	2.11
19	231	525.46	16.17	2.13	2.13
25	305	526.20	14.80	2.10	2.07
35	314	524.69	15.34	2.11	2.00
45	286	521.76	14.48	2.07	1.96
55	241	517.77	14.36	2.07	1.96
65	224	512.26	14.09	1.97	1.93
75	209	506.82	11.71	1.91	1.94
85	189	499.94	11.21	1.86	1.96

WJ V Nonverbal Cognitive Ability Index Construct Validity

This study examines the construct and predictive validity of the WJ V NVCAI as a psychoeducational assessment using WJ V norming data. Specifically, it aims to evaluate how well the NVCAI's internal structure aligns with the CHC theoretical model using confirmatory factor analysis (CFA) within the norming sample, providing evidence of construct validity (Lane & Marion, 2025; Kane, 2001, 2006). Analyses focused on school-age participants (ages 6–19; $N = 3,403$) with NVCAI scores. All analyses were conducted using IBM® SPSS® Amos 31 (IBM Corp., 2025).

Results

To evaluate the construct validity of the WJ V NVCAI, CFA was conducted with the norming sample of school-age participants (ages 6–19). A hypothesized three-factor model representing NV Fluid Reasoning (NV Gf), NV Visual Processing (NV Gv), and NV Processing Speed (NV Gs) was specified and tested. Maximum likelihood estimation produced the following fit indices: GFI = .957, AGFI = .892, CFI = .943, RMSEA = .117, and SRMR = .060 (see Table 11). Although the GFI, CFI, and SRMR suggested acceptable fit, the RMSEA value of .117 exceeded the conventional threshold of .08 and should be interpreted cautiously, particularly given the large sample size and the known sensitivity of RMSEA to model characteristics such as degrees of freedom. As shown in Figure 3, all seven tests loaded significantly onto their factors, with strong higher-order loadings supporting the proposed CHC-based structure.

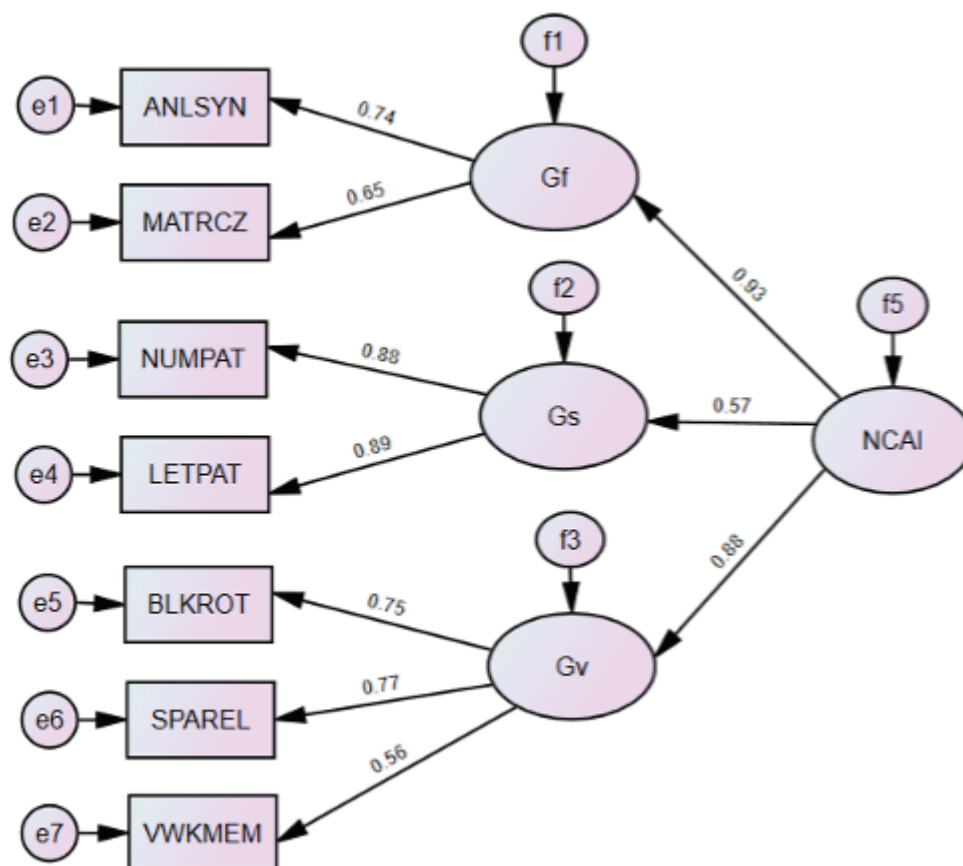
Table 11.
Comparison of CFA Fit
Statistics for the NVCAI for
Ages 6–19

	<i>df</i>	<i>X</i> ¹	GFI	AGFI	CFI	RMSEA	SRMR
ML	11	526.804	0.957	0.892	0.943	0.117 [0.109, 0.126]	0.060
SFLS	11	253.174	0.990	0.973	—	—	0.052

Notes. ML = maximum likelihood estimation algorithm, SFLS = scale-free least squares estimation algorithm, GFI = goodness-of-fit index, AGFI = adjusted goodness-of-fit index, CFI = comparative fit index, RMSEA = root mean square error of approximation, SRMR = standardized root mean square residual. Acceptable model fit was defined as GFI ≥ .90, AGFI ≥ .90, CFI ≥ .90 (good fit: ≥ .95), RMSEA between .05 and .08 (Browne & Cudeck, 1993), and SRMR < .08 (Hu & Bentler, 1999).

¹ Given multicollinearity, non-normality, and model complexity in the WJ V data, ML-based fit indices may be biased and should be used only for relative model comparisons. Therefore, SFLS estimation was employed as a complementary approach, as it does not assume multivariate normality (LaForte et al., 2025).

Figure 3.
Construct Validity of NVCAI
for Ages 6–19



Clinical Implications and Interpretive Guidance

The WJ V Nonverbal is intended to serve as a complementary tool within a comprehensive psychoeducational evaluation, not as a standalone measure of general intelligence. Practitioners should consider its use when there is reason to believe that the linguistic demands of standard cognitive assessment may introduce construct-irrelevant variance into test performance; that is, when language proficiency, expressive communication ability, or linguistic background may affect scores in ways that do not reflect the examinee's underlying cognitive functioning.

The cluster is particularly well-suited for use with English learners, students with developmental language disorders, students with speech-language impairments, and students with hearing impairments. For these populations, the language-minimized administration procedures may provide a more valid estimate of cognitive functioning than standard verbal administration alone. The cluster may also support more equitable evaluation of students from culturally and linguistically diverse backgrounds for whom the language demands of traditional cognitive batteries may not reflect actual cognitive capacity.

Practitioners should exercise caution, however, in assuming that nonverbal administration eliminates all sources of cultural or linguistic influence on performance. The WJ V Nonverbal is best understood as a language-minimized rather than a language-free measure. Visual and figural task formats are not culturally neutral, and performance may still be influenced by factors such as familiarity with testing materials, prior exposure to visually based academic tasks, and cultural variation in approaches to problem solving. The index is therefore most appropriately used as one component of a broader evaluation that incorporates multiple data sources, including developmental and educational history, language background information, observations of examinee behavior during testing, and other relevant assessment data.

The WJ V Nonverbal is not intended to replace standard cognitive assessment for examinees who do not present with language-based barriers to performance. For general education students without identified language concerns, standard administration of the WJ V COG provides a more comprehensive profile of cognitive functioning and should remain the primary assessment approach.

Verbal and Nonverbal Subtests Should Not be Combined in Composite Scores

Verbal and nonverbal subtests should *not* be combined when calculating the WJ V COG composite scores or the GIA. For example, an examinee should *not* complete Matrices, Spatial Relations, and Number-Pattern Matching in a nonverbal format and complete Oral Vocabulary, Story Recall, Semantic Word Retrieval, Verbal Attention, and Verbal Analogies in a verbal format, with all scores combined into a single GIA.

The WJ V Nonverbal cognitive tests were developed in part to provide a more accurate assessment of examinees with limited language skills. For these examinees, verbal subtests may yield lower scores that do not reflect their true cognitive abilities. If verbal and nonverbal scores were combined into a single composite, the GIA could be lowered by verbal subtest performance that is not valid for that examinee. As a result, the GIA may not provide a valid estimate of *g* for that individual, regardless of how well the nonverbal subtests perform.

This is the clinical situation in which the NVCAI serves as the more appropriate primary index. The concern is not that the NVCAI overestimates ability, but that the GIA may underestimate ability when verbal test performance is affected by language limitations.

Interpretive Cautions

Several considerations should guide interpretation of WJ V Nonverbal scores within a comprehensive evaluation framework.

First, scores from the WJ V Nonverbal should be interpreted considering the examinee's full linguistic and cultural profile. A higher score on the WJ V Nonverbal relative to verbally mediated cognitive measures may suggest that language demands are affecting performance on standard administration, but this pattern should be interpreted cautiously and in conjunction with other evaluation data rather than as a definitive indicator of language interference.

Second, practitioners should pay attention to behavioral observations during nonverbal administration. The WJ V Nonverbal relies entirely on video-based directions and gestural communication, a format that may be unfamiliar to some examinees. When an examinee appears uncertain about task expectations or requires additional orientation, their observed performance may not fully reflect their underlying cognitive ability. Therefore, it is important to note the examinee's apparent comfort with and understanding of the nonverbal procedures as part of valid score interpretation.

Third, the WJ V Nonverbal measures Visual Working Memory and a specific subset of CHC broad abilities (Fluid Reasoning, Visual Processing, and Processing Speed), and it should not be interpreted as a comprehensive index of overall cognitive functioning. Practitioners should avoid overinterpreting the WJ V NVCAI score as a global measure of intelligence and instead should situate it within the broader pattern of cognitive performance obtained across the full evaluation.

Finally, as with all standardized assessments, scores should never be interpreted in isolation. The WJ V Nonverbal is one source of data within a comprehensive evaluation process that considers the examinee's full developmental, educational, linguistic, and cultural background. Interpretation should always reflect the practitioner's professional judgment and knowledge of the individual examinee.

References

- Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Sage.
- Cormier, D. C., McGrew, K. S., & Evans, J. J. (2011). Quantifying the “degree of linguistic demand” in spoken intelligence test directions. *Journal of Psychoeducational Assessment*, 29(6), 515–533.
- de Brey, C., Musu, L., McFarland, J., Wilkinson-Flicker, S., Diliberti, M., Zhang, A., Branstetter, C., and Wang, X. (2019). *Status and Trends in the Education of Racial and Ethnic Groups 2018* (NCES 2019-038). U.S. Department of Education. National Center for Education Statistics. Retrieved August 10, 2026, from <https://nces.ed.gov/pubsearch/>
- Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1–55.
- IBM Corp. (2025). *IBM SPSS Amos* (Version 31.0) [Computer software]. IBM. <https://www.ibm.com/products/structural-equation-modeling-sem>
- Johnson, E. S. (2000). The effects of accommodations on performance assessments. *Remedial and Special Education*, 21(5), 261–267.
- Kane, M. T. (2001). Current concerns in validity theory. *Journal of Educational Measurement*, 38, 319–342.
- Kane, M. T. (2006). Validation. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 17–64). Praeger.
- Kaufman, A. S., & Kaufman, N. L. (2018). *Kaufman Assessment Battery for Children, Second Edition Normative Update*. NCS Pearson.
- Kaufman, A. S., Kaufman, N. L., Drozdick, L. W., & Morrison, J. (2018). Manual Supplement. *Kaufman Assessment Battery for Children Second Edition Normative Update*. NCS Pearson.
- LaForte, E. M., Dailey, D., & McGrew, K. S. (2025). Technical Manual. *Woodcock-Johnson V*. Riverside Assessments, LLC.
- Lane, S., & Marion, S. F. (2025). Validity and validation. In L. L. Cook & M. J. Pitoniak (Eds.), *Educational measurement* (5th ed., pp. 191–276). Oxford University Press.
- McCallum, R. S. (2017). Context for nonverbal assessment of intelligence and related abilities. In R. S. McCallum (Ed.), *Handbook of Nonverbal Assessment* (2nd ed., pp. 3–19). Springer.
- R Core Team. (2025). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Sireci, S. G., & Faulkner-Bond, M. (2015). Promoting validity in the assessment of English learners. *Review of Research in Education*, 39(1), 215–252.
- Thurlow, M. L., & Kopriva, R. J. (2015). Advancing accessibility and accommodations in content assessments for students with disabilities and English learners. *Review of Research in Education*, 39(1), 331–369.



Riverside
Insights®

800.323.9540

https://www.riversideinsights.com/woodcock_johnson_v