WISC®–V Coding and Symbol Search in Digital Format: Reliability, Validity, Special Group Studies, and Interpretation

Q-interactive Technical Report 12

Susan Engi Raiford, PhD
Ou Zhang, PhD
Lisa Whipple Drozdick, PhD
Kristen Getz, MA, CCC-SLP
Dustin Wahlstrom, PhD
Amy Gabel, PhD
James A. Holdnack, PhD
Mark Daniel, PhD

April, 2016
(Updated: August 2016)

Pearson

Abstract

This technical report provides information about the adaptation of the WISC–V Coding and Symbol Search subtests into digital format for Q-interactive™, Pearson's digital test administration and scoring platform. Administering, presenting stimuli, responding, and scoring of these subtests can all now be accomplished in a digital format within the Q-interactive platform beginning with their release in April 2016. No components external to Q-interactive are required.

In digital format, these subtests involve substantive changes to the examinee–tablet interactions, including onscreen touch responses, scrolling stimuli, and the elimination of writing requirements and self-corrected responses. Therefore, new data were collected and additional evidence of reliability and validity are provided in this technical report.

Introduction

In the initial phase of adapting published tests for the Q-interactive platform, the primary goal was to maintain raw-score equivalence between the subtests in both paper and digital administration and scoring formats. This goal was facilitated by minimizing effects of examinee–tablet interaction and assessment in the digital environment through maintaining use of external manipulatives. Equivalence of the paper and digital formats of the WISC–V was demonstrated by Daniel, Wahlstrom, and Zhang (2014), which allowed the norms, reliability, and validity information gathered for the WISC–V paper format to be applied to the digital format. For the September 2014 WISC–V release on Q-interactive, administration and scoring of all Processing Speed subtests involved using a paper response booklet to present stimuli and obtain examinee responses, a hard copy key or template to score the tasks manually, and the practitioner device (i.e., the iPad used by the practitioner) to enter the derived information into Q-interactive so scaled scores could be determined.

Design and development work on the Processing Speed subtests continued from September 2014 to February 2016. The continuing work culminated in the April 2016 release of the Coding and Symbol Search subtests for which administering, presenting stimuli, responding, and scoring all are accomplished in a digital format within the Q-interactive platform, eliminating the need for any external components. Scoring is automated and requires no effort on the part of the practitioner. Cancellation could not be adapted for a digital format, so it remains in its paper form and can be used on Q-interactive with a response booklet.

From the beginning of the WISC–V project, the goal was to establish a scaling relationship between the paper and digital formats of the Coding and Symbol Search subtests so practitioners could have confidence that either format would produce similar results for clinical use. Thus, prior to establishing the scaling relationship, the two formats of both subtests were studied to determine if they measure the same construct and have similar psychometric properties.

Coding and Symbol Search in digital format involve substantive changes to the examinee–tablet interactions relative to the paper format, including onscreen touch responses, scrolling stimuli, and the elimination of writing requirements and self-corrected responses. Therefore, new data were collected and additional evidence of reliability and validity for score interpretation were derived for this technical report. The research procedures, standardization, and scaling technique are described, and additional evidence of reliability and validity is provided.

Research, Standardization, and Scaling Procedures

The research program leading to the publication of Coding and Symbol Search in digital format was an iterative process that unfolded over a four-year period, with each stage of development leading to further refinements of the subtests. The *Standards for Educational and Psychological Testing* (*Standards*; American Educational Research Association [AERA], American Psychological Association [APA], & National Council on Measurement in Education [NCME], 2014) served as a primary resource throughout the project.

Major Research Stages

Conceptual Development Stage

The initial goal was to adapt the Processing Speed subtests for digital format without requiring use of paper response booklets while maintaining close similarity to the paper format. The research team at Pearson (described in chapter 3 of the WISC–V Technical and Interpretive Manual) consulted user interface designers, experts in cognitive ability testing, and the established literature on computerized testing to develop the initial concepts.

Pilot Stage

Pilot 1

The initial versions of Coding, Symbol Search, and Cancellation in digital format were established for the first pilot, which was conducted concurrently with the WISC–V standardization stage. The pilot 1 versions were kept as close to the paper format as possible without requiring pencil or paper. For example, each "page" of items (i.e., 9 items for Coding and 10 items for Symbol Search) required the child to touch an arrow to advance to the next page, so self-corrections could be accepted on the last items for each "page." The child used a stylus to respond for Coding. Symbol Search and Cancellation were programmed to allow touch responses, and Cancellation was split into four quadrants that were completed in four separate 15-second intervals. The sample for the first pilot was the same as the one described in Daniel et al. (2014).

It was assumed that the standardization paper format and pilot 1 digital format would not be raw-score equivalent because of the difference in response mode, but that they would be sufficiently correlated that the data from each format could be subjected to equating procedures. Results indicated, however, that the correspondence between the WISC–V standardization paper format and pilot 1 digital format was insufficient to support the use of equating procedures from both empirical and response process perspectives. Specifically, several issues arose. Empirically, some of the correlations did not reach a threshold of .7, and some were much lower (e.g., the correlation between the two formats of Coding for ages 6–7 was less than .4). In addition, a video review of live administrations conducted to study the response processes of the pilot 1 digital format revealed that requiring the examinee to touch an arrow at the bottom of each "page" of the Coding and Symbol Search items seemed to invoke selective attention and cognitive flexibility. This issue disrupted the flow of the perceptual speed task requirements. During the video review, the research team also noted children visibly paused and reoriented themselves when new "pages" of items suddenly appeared after touching the arrow. This effectappeared particularly pronounced in younger children (i.e., ages 6–7). In addition, despite the creation of inactive areas on the iPad, younger children tended to drop their wrist on the screen while writing with the stylus. This sometimes deactivated the response area in which they were writing and thus hindered the tablet's acceptance of responses.

Minipilot 1

The instructions of the pilot 1 versions of the subtests in digital format were adjusted to provide more teaching in an attempt to train out the pause and reorientation effect. Tablet cases and larger styli were acquired for use with young children ages 6–7. A minipilot (i.e., minipilot 1) with 30 examinees was conducted using the expanded instructions and the new cases (to lift the child's hand to prevent screen deactivation) and styli (to stabilize young children's grip and prevent their hands from slipping down and deactivating the screen). All administrations were captured on video. Input from examiners and review of the minipilot 1 administration videos suggested the pause and reorient effect remained despite the additional teaching. Furthermore, the case lifted children's hands into an awkward position, and the larger stylus resulted in an odd grip in some children.

Experts in user interface design for children reviewed the pilot 1 and minipilot 1 versions of the subtests. Consultation indicated that the reorienting effect in response to new stimuli that appear without a transition is a known phenomenon in the child media industry. Designs with smoother transitions (i.e., using shuffling stimuli) were recommended to prevent reorienting effects. The user interface design experts also confirmed that a stylus should be eliminated for these types of tasks for the reasons discovered in pilot 1 and minipilot 1. The user interface experts did not have additional suggestions for adapting Cancellation, so the subtest remained in its paper form and can be used on Q-interactive with a response booklet.

Minipilot 2

The experts in user interface design for children, as well as user interface designers internal to Pearson, collaborated with the research team to generate three prototypes of the Coding and Symbol Search subtests that applied the suggested concepts and eliminated the issues discovered in the previous editions. The prototypes were subjected to an internal expert review. A second minipilot (i.e., minipilot 2) using the prototypes was conducted with a sample of convenience (N=10), and all administrations were again captured on video and reviewed by the research team.

An advisory panel was formed, consisting of practitioners who were currently using Q-interactive to conduct clinical assessment and of psychologists who are nationally recognized experts in child clinical psychology, neuropsychology, and learning disabilities. The panel reviewed the new prototypes and provided input relevant to user friendliness; examinee reactions, behaviors, and observations; response processes; and construct measured. Results of semistructured surveys and interactive interviews suggested the prototypes appeared to measure processing speed, evoked the expected response processes, and avoided the drawbacks of prior versions.

Pilot 2

Following the minipilots, video review, and panel review, the collective feedback resulted in the conceptualization of a hybrid prototype of two of the three Coding versions, and one of the three prototyped versions of Symbol Search was selected. These selected prototypes were subjected to further development (e.g., full sets of items were built, data capture was enabled) and were carried forward for a second pilot (i.e., pilot 2).

Pilot 2 included 100 examinees (n = 50 for ages 6–7 and n = 50 for ages 8–9) without clinical conditions. A few children with intellectual disability were also tested to ensure instructions were clear and that they could comprehend the task. Approximately 10% of administrations with nonclinical children, and all administrations with children with intellectual disability, were videotaped. Several members of the research team participated as examiners. The goals of pilot 2 were to ensure confidence in the correlations of the paper and digital subtests, to examine the response processes in younger children to ensure the enhancements had resulted in improvement and eliminated the previous issues, and to examine the correlation of the subtests in digital format with one another. Results indicated the correlations of the paper and digital format of each subtest for both Form A and Form B exceeded .70. In addition, video review, interviews, and examiner notes indicated the response processes were as expected. The correlation of Coding and Symbol Search in digital format exceeded .5, as expected.

Pilot 3

Following pilot 2, the subtests were formally placed on Q-interactive to undergo a third pilot. For pilot 3, children without clinical conditions were randomly assigned to the digital format or to a waitlist condition, in anticipation of the counterbalanced research design planned for the standardization stage. A total of 70 examinees (n = 35 for ages 6–7 and n = 35 for ages 8–9) without clinical conditions was tested for pilot 3. The samples for both age bands were stratified to match census proportions within 4% of target, and gender within each age band was roughly equally distributed.

The goal of pilot 3 was to examine the distribution of scores for Coding and Symbol Search in digital format, the intercorrelations of Coding and Symbol Search in digital and in paper format, and the correlations of Coding and Symbol Search in digital format with the remainder of the 10 primary subtests. Concurrently, several children from special groups (i.e., intellectual giftedness, intellectual disability, attention-deficit/hyperactivity disorder, autism spectrum disorder) were also tested to examine specific research questions pertaining to the interaction of children with special conditions with new features of the subtests. For example, children with intellectual giftedness were tested and videotaped to examine their reactions to having no option to self-correct responses, and children with autism spectrum disorder were tested and videotaped to examine the potential impact of the scrolling stimuli on children with sensory issues. Results indicated that Coding and Symbol Search in digital format related to other subtests as expected, and that children with clinical conditions interacted with the digital subtests in the manner expected without adverse reactions to new features. In addition, video review, interviews, and examiner notes indicated the subtests in digital format were functioning as expected.

Standardization Stage

Based on the quality of results obtained with the pilot 3 version, the same versions of the subtests were used to collect data for the standardization stage. Because no changes were needed, the pilot 3 data were placed in the appropriate standardization studies.

The standardization stage focused on deriving standard scores and providing evidence of reliability, validity, and clinical utility for the final versions of the subtests in digital format. Data were obtained from a stratified sample of 329 children ages 6:0–16:11 who were administered Coding and Symbol Search in both digital and paper formats (i.e., the scaling sample). Coding and Symbol Search in digital format were inserted in the standard administration order (see Table 2.4 in the WISC–V Administration and Scoring Manual) with the remaining eight primary subtests. For children in the scaling sample, Coding and Symbol Search in paper format were given after the 10 primary subtests were completed. Samples of children from special groups were administered the same battery using the same procedure. A test-retest study and data for a counterbalanced correlational study with the paper format of the subtests were also collected.

Final Assembly and Evaluation Stage

In the final version, all Coding and Symbol Search subtest instructions to the child were maintained in identical form as standardization. Although the wording of instructions to the examiner may have been slightly modified to improve clarity, no modifications were made to alter the standardized administration procedures for the subtests. No changes were made to any stimuli.

Standardization

Locating and Testing the Samples

The participation criteria for the nonclinical studies and the special group studies were identical to those used for the normative and special group studies conducted for the WISC–V standardization phase and described in the WISC–V Technical and Interpretive Manual. The examiners and scorers were trained and data were collected and handled in accordance with the same procedures described in chapter 3 of the WISC–V Technical and Interpretive Manual.

Description of the Scaling Sample

The scaling sample was collected from July 2015 through January 2016 and is representative of the U.S. English-speaking population of children ages 6:0–16:11. A stratified sampling plan ensured that the normative sample included representative proportions of children according to selected demographic variables. An analysis of data gathered in 2013 by the U.S. Bureau of the Census provided the basis for stratification along the following variables within each age group: education level, gender, race/ethnicity, and geographic region. The following paragraphs present the characteristics of the nonclinical sample.

Age. The sample included 329 children divided into 11 age groups: 6:0–6:11, 7:0–7:11, 8:0–8:11, 9:0–9:11, 10:0–10:11, 11:0–11:11, 12:0–12:11, 13:0–13:11, 14:0–14:11,15:0–15:11, and 16:0–16:11. Because Coding and Symbol Search have separate forms for ages 6–7, the first two age groups were more heavily represented than the remaining age groups, with n = 45 and 49, respectively. The remaining age groups each contained approximately 20–30 participants.

Education Level. The sample was stratified according to four education levels based on the number of years of school completed by the parent(s). If the child resided with only one parent or guardian, the education level of that parent or guardian was assigned. If the child resided with both parents or with two guardians, the average of both individuals' education levels was used, with partial levels rounded up to the next highest level. The four education levels were defined as follows: 12 years or less completed with no high school diploma or equivalent, high school diploma or equivalent, 13–15 years (some college or associate's degree), and ≥16 years (college or graduate degree) of formal education.

Gender. The sample was not based on census percentages but consisted of a roughly equal number of female and male children in each age group.

Race/Ethnicity. For each age group in the normative sample, African Americans, Asians, Hispanics, Whites, and other racial/ethnic groups were represented.

Region. The United States was divided into the four major geographic regions specified by the census reports (see Figure 3.1 in the WISC–V Technical and Interpretive Manual): Midwest (MW), Northeast (NE), South (S), and West (W). Effort was made to ensure that each region was represented within each age group in the sample.

Representativeness of the Sample

Tables 1–2 present detailed demographic information for the scaling sample and the U.S. population according to age group, sex, race/ethnicity, parent education level, and geographic region.

Table 1. Percentages of the Scaling Sample and U.S. Population by Age Group, Parent Education Level, and Gender

Age Group		Parent Education Level				Gender	
		1	2	3	4	Female	Male
6	U.S. population	10.1	20.9	35.4	33.6	50.0	50.0
	Scaling sample	15.6	17.8	33.3	33.3	51.1	48.9
7	U.S. population	10.4	20.5	35.1	34.0	50.0	50.0
	Scaling sample	10.2	24.5	30.6	34.7	57.1	42.9
8	U.S. population	10.8	20.1	34.7	34.4	50.0	50.0
	Scaling sample	20.0	16.0	28.0	36.0	56.0	44.0
9	U.S. population	11.0	20.3	35.2	33.6	50.0	50.0
	Scaling sample	16.7	20.0	33.3	30.0	46.7	53.3
10	U.S. population	10.7	20.2	35.3	33.7	50.0	50.0
	Scaling sample	18.5	18.5	33.3	29.6	51.9	48.1
11	U.S. population	11.2	20.7	35.5	32.7	50.0	50.0
	Scaling sample	8.0	20.0	36.0	36.0	56.0	44.0
12	U.S. population	11.5	20.9	35.3	32.3	50.0	50.0
	Scaling sample	7.7	23.1	34.6	34.6	53.8	46.2
13	U.S. population	11.6	20.8	35.6	32.0	50.0	50.0
	Scaling sample	15.4	15.4	30.8	38.5	57.7	42.3
14	U.S. population	11.2	21.3	35.5	32.0	50.0	50.0
	Scaling sample	5.3	26.3	42.1	26.3	47.4	52.6
15	U.S. population	11.6	21.8	34.8	31.8	50.0	50.0
	Scaling sample	14.8	25.9	29.6	29.6	48.1	51.9
16	U.S. population	11.7	21.4	35.9	31.0	50.0	50.0
	Scaling sample	16.7	20.0	30.0	33.3	46.7	53.3

Table 2. Percentages of the Scaling Sample by Race/Ethnicity and Geographic Region

	Race/Ethnicity					Geographic region			
Age Group	African American	Asian	Hispanic	Other	White	Midwest	Northeast	South	West
6	13.3	—	15.6	13.3	57.8	6.7	31.1	44.4	17.8
7	14.3	2.0	26.5	6.1	51.0	10.2	8.2	55.1	26.5
8	—	—	32.0	—	68.0	8.0	20.0	52.0	20.0
9	10.0	3.3	36.7	10.0	40.0	16.7	13.3	50.0	20.0
10	—	—	44.4	11.1	44.4	7.4	11.1	48.1	33.3
11	4.0	—	32.0	4.0	60.0	12.0	16.0	60.0	12.0
12	—	7.7	23.1	7.7	61.5	3.8	11.5	30.8	53.8
13	7.7	—	19.2	7.7	65.4	19.2	11.5	46.2	23.1
14	10.5	5.3	21.1	—	63.2	21.1	10.5	42.1	26.3
15	14.8	—	22.2	11.1	51.9	7.4	11.1	63.0	18.5
16	16.7	—	16.7	6.7	60.0	16.7	16.7	50.0	16.7

Scaling Procedures

Prerequisites

Before applying equating methods, Dorans (2004) indicates a few requirements should be met. First, adequate reliability is necessary to ensure that results are informative. The test-retest coefficient was used because the split-half coefficient is not a proper reliability estimate for Coding or Symbol Search. Second, scores that have been placed on a common metric should only be equated if they measure the same construct with equal reliability and are related in the same way across different subpopulations and forms. Third, the correlation between the raw scores of the two forms must be high.

Inferential Scaling Method

There are multiple observed score equating methods available. The method used for the Coding and Symbol Search subtests is the inferential scaling method (Zhu & Chen, 2011) which involves transforming raw scores from two test forms into scores on the same standard scale to establish a relationship. According to Wilkins, Rolfhus, Weiss, and Zhu (2005), inferential scaling has several advantages over other traditional methods. First, it is useful with smaller sample sizes because it capitalizes on information across different age groups to develop estimates of the population raw score distributions for each of the age groups. That is, it maximizes the use of all available information so that the raw-score-to-scaled-score relationship for each age group is modified using information from the entire sample. Second, it involves expert knowledge and judgment to help determine the fit of a polynomial curve to the data. This improves the quality of scales by reducing sampling error. Third, it uses information about the population distribution in the estimation process. Fourth, it is based on smoothed and estimated population data. Finally, it allows any necessary hand smoothing, both within and across age groups, to ensure that connections between scores are established properly and are within tolerances relative to the age group and overall sample requirement. Based upon these strengths, inferential scaling is designed to provide more accurate results than other methods using similar sample sizes.

Results

As indicated by the results of the reliability and validity studies discussed later in this report, the data met the prerequisite requirements for equating methods. Based upon these favorable results, the inferential scaling method was selected and applied to the data. The raw scores from the digital format were subjected to inferential scaling procedures to create scaled scores. Consequently, the raw scores derived from the digital and paper format were on the same metric (i.e., scaled scores). The scaled scores then were subjected to the analyses presented later in this report, which demonstrate that there are no meaningful differences between the digital scaled scores and the paper scaled scores obtained from inferential scaling.

When the digital format is administered, the raw score and the scaled score that is output on the subtest summary card and the results tab in Assess (Q-interactive's tablet app) and on Central (Q-interactive's online web-based portal) pertain to the digital format. Therefore, no hand-scoring or adjustment is needed.

Evidence of Reliability

The evidence of test-retest stability for the Coding and Symbol Search scaled scores was obtained by administering the digital format of the subtests on two occasions. Table 3 presents the demographic characteristics and testing interval statistics for the test-retest reliability sample.

Test-retest stability was estimated for two age bands (i.e., 6–7, 8–16) and for all ages. The mean Coding and Symbol Search subtest scores and their *SD*s for all ages and for each of the two age bands are presented in Table 4. The average stability coefficients for all ages were calculated using Fisher's *z* transformation. The table also reports the standard differences (i.e., effect sizes) between the first and second testing and the correlation coefficients corrected for the theoretical standard deviation (i.e., 3.0). The standard difference was calculated using the mean score difference between the two testings, divided by the square root of the pooled variance (Cohen, 1988).

As the data in Table 4 indicate, Coding has adequate stability and Symbol Search has good stability across time for all ages. As the data also indicate, the mean retest scores for both subtests are higher than the mean scores from the first testing. These results are consistent with those reported for the paper format (Wechsler, 2014).

Table 3. Demographic Characteristics of Test-Retest Reliability Sample

Sample	Ages 6–7	Ages 8–16	All Ages
N	33	35	68
Age			
Mean	7.0	12.4	9.8
SD	0.6	3.2	3.6
Range	6-7	8-16	6-16
Testing Interval			
Interval mean	25.8	24.8	25.3
Interval range	14–72	14–52	14–72
Parent Education			
0–12 years of school, no diploma	6.1	11.4	8.8
High school diploma or equivalent	6.1	14.3	10.3
Some college or technical school, associate's degree	33.3	40.0	36.8
Bachelor's degree	54.5	34.3	44.1
Race/Ethnicity			
African American	12.1	11.4	11.8
Asian	3.0	5.7	4.4
Hispanic	21.2	28.6	25.0
Other	3.0	11.4	7.4
White	60.6	42.9	51.5
Region			
Midwest	21.2	8.6	14.7
Northeast	3.0	8.6	5.9
South	57.6	57.1	57.4
West	18.2	25.7	22.1
Sex			
Female	51.5	37.1	44.1
Male	48.5	62.9	55.9

Note. Except for sample size, testing interval, and age, data are reported as percentages. Total percentage may not add up to 100 due to rounding.

Table 4. Stability Coefficients of Coding and Symbol Search (Digital Format)

Overall	First testing		Second testing			Corrected	Standard
	Mean	*SD*	Mean	*SD*	*r*	*r*	Difference
Coding Scaled Score	11.0	3.5	12.0	3.4	.80	.75	.29
Symbol Search Scaled Score	11.6	2.9	13.3	3.4	.78	.80	.54
Age 6–7	**First testing**		**Second testing**			**Corrected**	**Standard**
	Mean	***SD***	**Mean**	***SD***	***r***	***r***	**Difference**
Coding Scaled Score	10.3	2.8	11.1	2.9	.74	.77	.28
Symbol Search Scaled Score	11.3	2.9	13.5	3	.77	.79	.75
Age 8–16	**First testing**		**Second testing**			**Corrected**	**Standard**
	Mean	***SD***	**Mean**	***SD***	***r***	***r***	**Difference**
Coding Scaled Score	11.6	4.0	12.9	3.6	.85	.73	.34
Symbol Search Scaled Score	11.9	2.8	13.1	3.8	.78	.81	.36

Evidence of Validity

Evidence Based on Internal Structure

Intercorrelations

Several a priori hypotheses were proposed for the intercorrelation study of Coding and Symbol Search in digital format with the other primary subtests and index scores. As with the paper version, it was assumed that all scores would show low to moderate correlations with one another, based on the assumption that all subtests measure a general intelligence factor (i.e., *g*). Second, it was expected that Coding and Symbol Search would have higher correlations with each other than with subtests from other cognitive domains. Third, it was anticipated that the eight remaining primary subtests would correlate more highly with each other than with Coding and Symbol Search because Processing Speed subtests typically have a weaker relationship with general intelligence relative to subtests from other cognitive domains.

Data were obtained from a sample of children who were administered Coding and Symbol Search in both the digital and paper formats, in counterbalanced order. Within each cell in a matrix that considered age, parent education level, and gender, half of the children were randomly assigned to be administered the digital format first or the paper format first. For each order, the first format of Coding and Symbol Search was inserted in the standard administration order (see Table 2.4 in the WISC–V Administration and Scoring Manual) with the remaining eight primary subtests, and the second format was given after the 10 primary subtests were completed. The demographic characteristics of the sample used for the intercorrelation study appear in Table 5. Overall, the sample has nearly equal representation of each age, and has nationally representative proportions of racial/ethnic groups. Hispanics are somewhat overrepresented relative to the general population.

The average intercorrelations of the subtest and composite scores for all ages (computed using Fisher's *z* transformation), and for the two age bands corresponding with the two forms of Coding and Symbol Search (i.e., 6–7 and 8–16) were calculated. The intercorrelations are presented in Table 6. The table includes the correlations of the subtests with the sums of scaled scores for each of the composites. The correlation between the sum of scaled scores for a composite and the scaled score for each contributing subtest was corrected by removing that subtest scaled score from the sum of the scaled scores to control for inflation of the values. The uncorrected coefficients appear below the diagonal and the corrected coefficients appear above the diagonal, in the shaded area.

Table 5. Demographics of Intercorrelation Study

N	651
Age	
Mean	10.9
SD	3.3
Range	6–16
Parent Education	
0–12 years of school, no diploma	12.6
High school diploma or equivalent	20.4
Some college or technical school, associate's degree	33.5
Bachelor's degree	33.5
Race/Ethnicity	
African American	11.8
Asian	1.4
Hispanic	24.4
Other	7.1
White	55.3
Region	
Midwest	12.3
Northeast	14.0
South	48.5
West	25.2
Sex	
Female	51.8
Male	48.2

Note. Except for sample size, testing interval, and age, data are reported as percentages. Total percentage may not add up to 100 due to rounding.

Table 6. Intercorrelations of Subtests and Composite Scores
Overall Sample

Subtest/ Composite Scores	SI	VC	BD	VP	MR	FW	DS	PS	CD	SS	VCI	VSI	FRI	WMI	PSI	FSIQ	GAI	NVI	CPI
SI											.54					.55	.56		
VC	.54										.54					.51	.52		
BD	.35	.35										.55				.52	.49	.56	
VP	.32	.33	.55									.55						.55	
MR	.37	.30	.45	.46									.37			.52	.51	.52	
FW	.34	.35	.31	.35	.37								.37			.47	.46	.45	
DS	.35	.32	.33	.32	.36	.29								.43		.49			.42
PS	.30	.31	.31	.32	.29	.32	.43							.43				.46	.45
CD	.18	.18	.23	.15	.18	.20	.26	.33							.53	.31		.31	.51
SS	.21	.23	.30	.25	.25	.22	.29	.28	.53						.53				.49
VCI	.88	.88	.39	.37	.37	.38	.39	.35	.20	.26									
VSI	.38	.39	.88	.89	.52	.37	.37	.36	.22	.31	.43								
FRI	.43	.38	.46	.49	.84	.82	.40	.38	.23	.29	.46	.54							
WMI	.39	.37	.38	.38	.39	.37	.83	.86	.35	.34	.43	.43	.45						
PSI	.22	.24	.31	.23	.25	.24	.32	.34	.87	.88	.27	.31	.31	.39					
FSIQ	.70	.67	.67	.55	.68	.63	.64	.50	.51	.46	.78	.69	.79	.68	.55				
GAI	.73	.71	.69	.56	.72	.66	.48	.44	.29	.36	.82	.71	.84	.54	.37	.95			
NVI	.48	.45	.72	.72	.70	.64	.50	.64	.53	.46	.53	.82	.81	.68	.57	.90	.85		
CPI	.36	.36	.42	.36	.38	.36	.67	.70	.75	.75	.41	.44	.45	.81	.85	.73	.54	.75	
Mean	9.9	9.9	10.2	10.1	10.5	10.2	10.2	10.9	10.2	10.5	19.8	20.4	20.7	21.1	20.7	71.1	50.8	62.0	41.8
SD	2.8	2.8	2.8	2.8	3.1	2.9	2.7	2.9	3.0	3.0	5.0	5.0	5.1	4.7	5.3	13.0	10.3	11.6	8.2
N	642	643	650	643	648	644	650	649	638	639	638	642	641	648	627	615	629	617	624

WISC–V abbreviations are: SI = Similarities, VC = Vocabulary, BD = Block Design, VP = Visual Puzzles, MR = Matrix Reasoning, FW = Figure Weights, DS = Digit Span, PS = Picture Span, CD = Coding, SS = Symbol Search, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, GAI = General Ability Index, NVI = Nonverbal Index, CPI = Cognitive Proficiency Index.

Table 6. Intercorrelations of Subtests and Composite Scores (*continued*)

Ages 6–7

Subtest/ Composite Scores	SI	VC	BD	VP	MR	FW	DS	PS	CD	SS	VCI	VSI	FRI	WMI	PSI	FSIQ	GAI	NVI	CPI
SI											.50					.54	.54		
VC	.50										.50					.46	.45		
BD	.36	.31										.52				.50	.48	.54	
VP	.30	.31	.52									.52						.52	
MR	.34	.26	.45	.45									.35			.50	.50	.52	
FW	.28	.22	.24	.23	.35								.35			.38	.37	.39	
DS	.36	.29	.33	.31	.31	.18								.39		.47			.49
PS	.31	.25	.33	.36	.29	.29	.39							.39				.49	.46
CD	.22	.24	.22	.16	.20	.24	.35	.37							.53	.38		.35	.55
SS	.27	.29	.30	.25	.32	.22	.40	.36	.53						.53				.56
VCI	.88	.86	.38	.35	.34	.28	.37	.33	.26	.32									
VSI	.38	.36	.85	.89	.52	.27	.37	.41	.22	.31	.42								
FRI	.38	.29	.43	.42	.84	.80	.31	.37	.27	.34	.38	.48							
WMI	.40	.32	.40	.41	.36	.29	.80	.86	.43	.45	.42	.47	.40						
PSI	.28	.31	.31	.24	.31	.27	.44	.41	.85	.89	.34	.31	.36	.50					
FSIQ	.70	.63	.66	.51	.67	.56	.63	.50	.56	.53	.77	.66	.75	.67	.62				
GAI	.73	.66	.68	.52	.72	.60	.44	.45	.34	.42	.80	.69	.81	.53	.44	.95			
NVI	.47	.40	.70	.71	.70	.59	.48	.67	.55	.50	.50	.81	.79	.70	.60	.88	.84		
CPI	.39	.37	.41	.37	.38	.32	.70	.71	.76	.79	.43	.44	.43	.85	.88	.74	.55	.75	
Mean	9.9	9.5	10.3	10.1	10.5	10.0	9.8	10.4	10.3	10.6	19.4	20.3	20.5	20.2	20.8	70.3	50.1	61.6	41.1
SD	3.0	2.8	2.8	3.3	3.1	2.8	2.6	3.0	2.8	3.2	5.0	5.3	4.8	4.6	5.3	12.5	9.9	11.6	8.6
N	178	179	182	179	181	180	182	182	180	179	177	179	179	182	177	174	176	174	177

WISC–V abbreviations are: SI = Similarities, VC = Vocabulary, BD = Block Design, VP = Visual Puzzles, MR = Matrix Reasoning, FW = Figure Weights, DS = Digit Span, PS = Picture Span, CD = Coding, SS = Symbol Search, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, GAI = General Ability Index, NVI = Nonverbal Index, CPI = Cognitive Proficiency Index.

Table 6. Intercorrelations of Subtests and Composite Scores (*continued*)

Ages 8–16

Subtest/ Composite Scores	SI	VC	BD	VP	MR	FW	DS	PS	CD	SS	VCI	VSI	FRI	WMI	PSI	FSIQ	GAI	NVI	CPI
SI											.58					.55	.57		
VC	.58										.58					.56	.59		
BD	.33	.38										.58				.53	.50	.58	
VP	.34	.34	.58									.58						.58	
MR	.39	.33	.44	.47									.39			.53	.51	.52	
FW	.39	.46	.38	.46	.39								.39			.55	.54	.50	
DS	.34	.35	.33	.32	.41	.40								.46		.50			.35
PS	.29	.36	.28	.28	.29	.35	.46							.46				.43	.43
CD	.13	.12	.24	.14	.15	.16	.16	.28							.53	.23		.27	.46
SS	.15	.17	.29	.24	.18	.21	.17	.20	.53						.53				.42
VCI	.88	.89	.39	.38	.40	.47	.40	.37	.14	.19									
VSI	.38	.41	.90	.88	.51	.47	.36	.31	.21	.30	.44								
FRI	.47	.47	.49	.55	.84	.83	.48	.38	.19	.24	.53	.59							
WMI	.37	.42	.36	.35	.41	.44	.85	.86	.26	.21	.44	.39	.50						
PSI	.16	.16	.31	.22	.19	.21	.19	.27	.88	.87	.19	.30	.25	.27					
FSIQ	.69	.70	.68	.58	.69	.69	.65	.50	.45	.38	.78	.71	.83	.68	.48				
GAI	.73	.75	.69	.60	.72	.72	.51	.42	.23	.29	.83	.73	.86	.55	.30	.95			
NVI	.49	.50	.73	.72	.70	.68	.51	.61	.51	.41	.55	.82	.83	.66	.53	.91	.86		
CPI	.33	.35	.42	.35	.37	.40	.63	.69	.74	.70	.39	.43	.46	.77	.82	.72	.52	.75	
Mean	9.9	10.1	10.2	10.2	10.5	10.2	10.4	11.0	10.1	10.5	20.0	20.4	20.7	21.4	20.6	71.5	51.0	62.2	42.1
SD	2.8	2.8	2.8	2.7	3.1	3.0	2.8	2.8	3.1	2.9	4.9	4.9	5.1	4.7	5.3	13.2	10.5	11.5	8.0
N	464	464	468	464	467	464	468	467	458	460	461	463	462	466	450	441	453	443	447

WISC–V abbreviations are: SI = Similarities, VC = Vocabulary, BD = Block Design, VP = Visual Puzzles, MR = Matrix Reasoning, FW = Figure Weights, DS = Digit Span, PS = Picture Span, CD = Coding, SS = Symbol Search, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, GAI = General Ability Index, NVI = Nonverbal Index, CPI = Cognitive Proficiency Index.

All intersubtest correlations are statistically significant, and all subtests correlate to some degree with each other. The pattern of intercorrelations is very similar to the patterns found for the WISC–V paper format (Wechsler, 2014) and other Wechsler intelligence tests (Wechsler, 2003, 2008, 2012). The correlations between the Processing Speed subtests and all other primary subtests are low, as anticipated. Of the primary subtests that do not contribute to the FSIQ, Symbol Search has the lowest correlation with the FSIQ. When the Coding correlation is corrected, it shares the lowest correlation with the FSIQ of all of the primary subtests. Similarly, the PSI is less correlated with the FSIQ than are the other primary index scores. Although the correlation values vary slightly, these patterns generally are repeated across the two age bands.

The consistency of the present findings with those obtained from a study of the WISC–V administered in paper format (Wechsler, 2014) indicates that Coding, Symbol Search, and the PSI are measuring similar constructs in both formats. These results provide evidence that when Coding and Symbol Search are administered in digital format, the WISC–V produces scores that are consistent with results obtained from administration of the paper format.

Confirmatory Factor-Analytic Studies

Models

The sample for the confirmatory factor-analytic studies was the same as the one described subsequently in this technical report for the paper-digital format equivalence study. This sample was also administered Coding and Symbol Search in paper format, in counterbalanced order. The demographic characteristics of the sample appear in Table 5. The WISC–V factor model corresponded to its theoretical design. A second-order factor model with five first-order factors (the five ability domains) and one second-order factor (general intelligence, or *g*) was supported for the WISC–V in paper format (Wechsler, 2014). The same model was anticipated for the WISC–V in digital format.

The first model in the series (Model 1) was a hierarchical model with general ability as a higher-order factor and first order factors representing Verbal Comprehension, Visual Spatial, Fluid Reasoning, Working Memory, and Processing Speed domains. The subtests from those domains loaded on their respective factors. Model 2 replaced the digital format of Coding and Symbol Search with the paper format of these subtests. If the model fit is not noticeably better for either format, the data are similarly explained by the model regardless of format, and construct equivalence of the digital and paper format subtests are supported. The change in fit is evaluated through a subjective comparison rather than statistical comparison of the two models' fit statistics because the models are not nested. The factor structure was analyzed for two age bands corresponding with the age bands for the two forms (i.e., 6–7 and 8–16) and overall (for the entire age range).

Results

Table 7 presents the fit statistics for the confirmatory factor analyses, based on the two age bands and the entire age range.

Table 7. Goodness-of-Fit Statistics for Confirmatory Factor Analyses

Model	X^2	*df*	*CFI*	*TLI*	*RMSEA*	*AIC*	*BIC*
Ages 6–7							
Model 1	29.6	25	.99	.93	.03	90	184
Model 2	31.4	25	.98	.92	.04	91	185
Ages 8–16							
Model 1	72.4	25	.96	.94	.07	132	254
Model 2	65.6	25	.97	.95	.06	126	247
All Ages							
Model 1	63.9	25	.98	.96	.05	124	256
Model 2	63.4	25	.98	.96	.05	123	255

Note. The chi-square values are weighted least squares from SAS® 9.3.

Model 1 includes Coding and Symbol Search in digital format. Model 2 includes Coding and Symbol Search in paper format.

The results shown in Table 7 indicate that both of the models have excellent fit. These results provide strong support that the subtests are measuring similar constructs whether in digital or paper format.

Figure 1 shows the subtest and factor loadings for Model 1 based on the analysis for all ages.

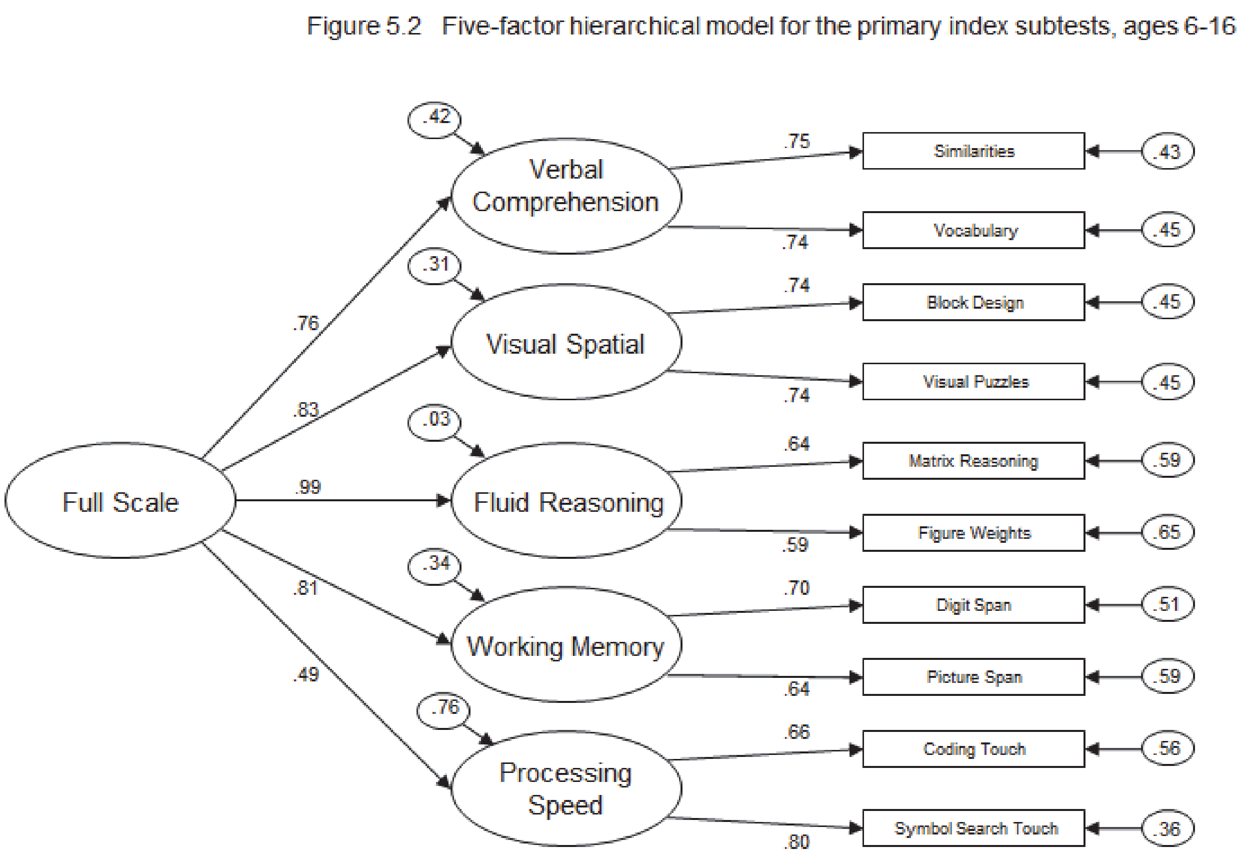

Figure 1. Factor Model for Primary Subtests in Digital Format, Ages 6–16

The subtest *g* loadings are presented in Table 8. The *g* loadings are similar for subtests within domains. The *g*-loadings differ slightly from those reported for the WISC–V in paper format because only the primary subtests were included, and loadings change depending on the combination of subtests placed in the model. The lowest loadings are for the Processing Speed subtests, similar to the paper format.

Table 8. Subtest *g* Loadings

Subtest	*g* Loading
Similarities	.57
Vocabulary	.56
Block Design	.61
Visual Puzzles	.61
Matrix Reasoning	.65
Figure Weights	.60
Digit Span	.57
Picture Span	.52
Coding	.32
Symbol Search	.39

Evidence Based on Relations With Other Variables

Paper-Digital Format Equivalence

The sample described for the intercorrelation study was also used to study paper-digital format equivalence. A test-retest design is appropriate when the response processes are unlikely to change substantially upon retest because the examinee does not learn solutions or new strategies for approaching the task or solving the problem. The demographic characteristics of the sample appear in Table 5. When a retest design is possible, it is powerful because examinees serve as their own controls.

For all Q-interactive equivalence studies, an effect size of 0.2 or smaller has been established by Pearson as the standard for equivalence. The effect size is the average amount of difference between scores on digital and paper format administrations, divided by the standard deviation of scores in the population. An effect size of 0.2 is slightly more than one-half of a scaled-score point on the commonly used subtest metric that has a mean of 10 and standard deviation of 3.

Table 9 reports the means, standard deviations, uncorrected and corrected correlations, and standard differences of Coding and Symbol Search subtest scores for each format. Given the close similarity of the demographic characteristics of the two format groups and the fact that children were randomly assigned to a format, one would not expect large or systematic differences in scores between the groups.

Table 9. Coding and Symbol Search Format Equivalence

	Paper				Digital								
	Raw Score		Scaled Score		Raw Score		Scaled Score		Raw Score		Scaled Score		
Subtest	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	r_{12}	*Corrected* r_{12}	r_{12}	*Corrected* r_{12}	Standard Difference
Coding	44.5	20.1	9.5	2.9	37.1	10.2	10.2	3.0	.87	.89	.63	.69	0.23
Symbol Search	28.4	9.7	10.9	3.1	29.5	9.3	10.5	3.0	.84	.85	.67	.68	-0.13

Note. Standard difference is always calculated by digital form standard score mean minus paper form standard score mean.

The raw score correlations are high and of the magnitude recommended for equating procedures (Dorans, 2004). The scaled score correlations are also high. Coding showed a format effect size that only slightly exceeded the 0.2 criterion, and Symbol Search showed a negligible effect size. The scaled scores used for this analysis are those that are applied for the digital format. Effects of this magnitude demonstrate that there are no meaningful differences between the obtained paper scaled scores and the obtained digital scaled scores. Taken together, these results provide strong evidence of equivalence of the scores derived from digital and paper formats of these subtests.

Special Group Studies

Evidence of a scale's validity, when applied to clinical and special groups, is crucial when its results are part of a comprehensive diagnostic assessment. Seven special group studies were conducted to determine if children in selected criterion groups with known characteristics perform as expected when Coding and Symbol Search are administered in digital format. Independent examiners and researchers collected the data for the special group studies. Candidates for these studies met the same criteria specified for these groups in Appendix A of the WISC–V Technical and Interpretive Manual.

It is important to note the limitations of these studies. The samples were not randomly selected but were recruited based on availability. Therefore, these studies may not be representative of performance of all children in the diagnostic category. Because data for each special group sample were collected in a variety of clinical settings, the diagnoses of children within the same special group might have been made on the basis of different criteria and procedures, particularly in the Specific Learning Disorders group. In addition, the sample sizes for some of the studies are small and cover only a portion of the WISC–V age range. Finally, only group performance is reported. For these reasons, the data from these samples are presented as examples and are not intended to be fully representative of these diagnostic groups. The purpose of the studies is to provide evidence that when Coding and Symbol Search are administered in digital format, children in these special groups obtain scores that are valid estimates of intellectual ability and score as expected.

Comparison groups were derived from the nonclinical scaling sample (previously described in this report). The comparison groups were matched to each special group according to age, sex, race/ethnicity, and parent education level. Cases were then randomly sampled such that the derived sample was matched to the clinical group on the constraining variables. The demographic characteristics of the special group study samples are presented in Table 10.

Tables 11–17 report the mean WISC–V performance of the special groups and their corresponding matched control groups. Coding, Symbol Search, and all other primary subtests were administered in digital format.

Table 10. Demographic Characteristics of Special Group Studies

	GT	IDMI	SLD-R	SLD-M	ADHD	ASD-L	MI
N	49	63	24	22	21	26	15
Age							
Mean	11.7	11.4	12.1	13.2	12.2	11.5	11.5
SD	3.2	3.3	2.9	2.6	3	2.8	3.7
Range	6–16	6–16	6–16	6–16	6–16	6–16	6–16
Parent Education							
0–12 years of school, no diploma	–	25.4	–	–	4.8	7.7	6.7
High school diploma or equivalent	2.0	31.7	12.5	40.9	19.0	11.5	26.7
Some college or technical school, associate's degree	6.1	27.0	54.2	45.5	9.5	42.3	20.0
Bachelor's degree	91.8	15.9	33.3	13.6	66.7	38.5	46.7
Race/Ethnicity							
African American	4.1	33.3	–	18.2	4.8	7.7	–
Asian	8.2	1.6	–	–	4.8	–	–
Hispanic	8.2	15.9	50.0	22.7	14.3	11.5	13.3
Other	10.2	–	4.2	4.5	4.8	7.7	6.7
White	69.4	49.2	45.8	54.5	71.4	73.1	80.0
Region							
Midwest	53.1	20.6	8.3	9.1	9.5	11.5	–
Northeast	2.0	1.6	12.5	–	–	3.8	20.0
South	18.4	73.0	70.8	81.8	42.9	26.9	73.3
West	26.5	4.8	8.3	9.1	47.6	57.7	6.7
Sex							
Female	42.9	41.3	41.7	45.5	19.0	26.9	20.0
Male	57.1	58.7	58.3	54.5	81.0	73.1	80.0

Note. Except for sample size and age, data are reported as percentages. Total percentage may not add up to 100 due to rounding.

Abbreviations are: GT = Intellectually Gifted, IDMI = Intellectual Disability-Mild Severity, SLD-R = Specific Learning Disorder-Reading, SLD-M = Specific Learning Disability-Mathematics, ADHD = Attention-Deficit/Hyperactivity Disorder, ASD-L = Autism Spectrum Disorder With Language Impairment, MI = Motor Impairment.

Children Identified as Intellectually Gifted

Children with intellectual giftedness typically show particular strengths in the areas of verbal comprehension, visual spatial ability, and fluid reasoning. Although their working memory and processing speed performance is generally higher than in the general population (Elliot, 2007; Kaufman & Kaufman, 2004; Wechsler, 2012, 2014), it typically is lower than their performance on verbal comprehension, visual spatial, and fluid reasoning domains (Raiford, Holdnack, Drozdick, & Zhang, 2014; Raiford, Weiss, Rolfhus, & Coalson, 2005; Rimm, Gilman, & Silverman, 2008; Rowe, Kingsley, & Thompson, 2010; Wechsler, 2012, 2014).

Raiford et al. (2014) demonstrated that very similar results are obtained with intellectually gifted children when using the WISC–V in digital format relative to paper format. Apart from this, there are few studies using digital instruments with intellectually gifted children. In general, a synthesis of the literature indicates that the use of digital technology in assessment and instruction is viewed positively by gifted children, and that digital technology can be utilized with gifted individuals to produce comparable or superior assessment and instruction results relative to traditional paper delivery (Periathiruvadi & Rinn, 2012). Studies with gifted children involving assessment of constructs related to intellectual ability, such as strategic thinking (Steiner, 2006) and self-regulation (Calero, García-Martín, Jiménez, Kazén, & Araque, 2007), indicate that assessment with gifted children produces comparable results in paper and digital formats.

The WISC–V in digital format was administered to a sample of children identified as intellectually gifted. The demographic characteristics of this sample appear in Table 10. Table 11 presents the mean subtest and composite scores for the Intellectually Gifted and matched control groups.

Table 11. Intellectually Gifted Compared to Matched Controls

Subtest/ Composite Score	Intellectually Gifted		Matched Control		Group Mean Comparison			
	Mean	*SD*	Mean	*SD*	Difference	*t* value	*p* value	Standard Difference[a]
SI	15.7	2.6	11.1	2.2	-4.60	-9.77	<.01	-1.91
VC	15.4	2.3	11.2	2.5	-4.22	-8.21	<.01	-1.76
BD	14.3	2.3	11.3	2.9	-2.96	-5.67	<.01	-1.13
VP	14.0	2.2	11.5	2.5	-2.50	-4.82	<.01	-1.06
MR	13.8	2.8	10.9	3.5	-2.88	-3.95	<.01	-0.91
FW	14.0	2.5	11.4	3.0	-2.66	-4.35	<.01	-0.96
DS	14.0	2.8	11.0	2.6	-3.00	-5.44	<.01	-1.11
PS	13.9	3.0	11.5	2.8	-2.47	-4.45	<.01	-0.85
CD	12.9	3.1	10.7	3.2	-2.13	-3.66	<.01	-0.68
SS	13.0	3.2	10.9	2.9	-2.10	-3.70	<.01	-0.69
VCI	131.0	13.1	106.1	10.3	-24.88	-9.73	<.01	-2.11
VSI	123.5	10.9	108.1	13.6	-15.43	-5.53	<.01	-1.25
FRI	123.3	11.2	106.3	16.2	-17.09	-5.02	<.01	-1.23
WMI	122.4	14.0	107.2	13.0	-15.22	-5.82	<.01	-1.13
PSI	116.8	16.8	104.7	15.4	-12.06	-4.21	<.01	-0.75
FSIQ	131.0	9.7	108.0	12.9	-23.02	-9.61	<.01	-2.02
NVI	127.4	10.6	108.9	13.8	-18.50	-6.40	<.01	-1.50
GAI	130.5	9.4	107.5	13.3	-23.07	-8.66	<.01	-2.00
CPI	123.5	15.2	107.5	13.3	-15.98	-6.29	<.01	-1.12

[a] The Standard Difference is the difference of the two test means divided by the square root of the pooled variance, computed using Cohen's (1996) Formula 10.4.

WISC–V abbreviations are: SI = Similarities, VC = Vocabulary, BD = Block Design, VP = Visual Puzzles, MR = Matrix Reasoning, FW = Figure Weights, DS = Digit Span, PS = Picture Span, CD = Coding, SS = Symbol Search, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, NVI = Nonverbal Index, GAI = General Ability Index, CPI = Cognitive Proficiency Index.

The results of this study are consistent with results of prior research involving the paper format of Coding and Symbol Search (Wechsler, 2003, 2014). All mean primary subtest and composite scores are significantly higher than those of the matched control group, and all differences have moderate to large effect sizes. The largest effects among primary subtests occur on Similarities and Vocabulary, and the smallest occur on Coding and Symbol Search. The highest mean subtest scores are on Similarities and Vocabulary, and the lowest occur on Coding and Symbol Search. Effect sizes for most index score differences are large. As is typically seen in samples of intellectually gifted children, the VCI has the largest effect size of all primary index scores. The PSI effect size is moderate and the smallest of all primary index scores.

Additional analysis indicates that of children identified as intellectually gifted, 85% receive WISC–V FSIQ scores of 120 points or higher, 86% have GAI scores of 120 points or higher, and 79% have NVI scores of 120 points or higher. In contrast, only 24%, 21%, and 32% of children in the matched control group achieve these respective scores.

The consistency of the present findings with those obtained from a study of the WISC–V administered in paper format (Wechsler, 2014) indicates that Coding, Symbol Search, and the PSI are measuring similar constructs in both formats. Taken together, these results indicates that when Coding and Symbol Search (and all of the remaining WISC–V primary subtests) are administered in digital format, the WISC–V produces scores that are useful in the assessment of intellectual giftedness and consistent with results obtained when administering the test in paper format. These results also add to the body of research that indicates children identified as intellectually gifted show lower performance on processing speed tasks than on those from other cognitive domains.

Children With Intellectual Disability

Many studies have been conducted to evaluate the performance of individuals with intellectual disability on previous versions of the Wechsler intelligence scales. In some studies, verbal comprehension and working memory performance is lower than visual spatial, fluid reasoning, and processing speed performance (Gordon, Duff, Davidson, & Whitaker, 2010). The smallest effect sizes for mean differences between intellectual disability and matched control groups are often on the processing speed measures. In addition, standard deviations of subtest and composite scores are generally smaller in intellectual disability groups than in the general population; however, this difference does not usually occur on Processing Speed subtests or the PSI (Nunes et al., 2012; Raiford et al., 2014; Wechsler, 2012, 2014).

The WISC–V in digital format was administered to a sample of children with intellectual disability-mild severity. The demographic characteristics of this sample appear in Table 10. Table 12 presents the mean subtest and composite scores for the Intellectual Disability and matched control groups.

Table 12. Intellectual Disability-Mild Compared to Matched Controls

Subtest/ Composite Score	Intellectual Disability-Mild Mean	Intellectual Disability-Mild *SD*	Matched Control Mean	Matched Control *SD*	Group Mean Comparison Difference	*t* value	*p* value	Standard Difference[a]
SI	3.5	2.3	9.4	3.1	5.89	12.7	<.01	2.16
VC	3.6	1.9	9.3	2.9	5.68	12.82	<.01	2.32
BD	4.5	2.4	9.5	2.7	4.97	11.54	<.01	1.95
VP	4.0	1.9	9.7	2.7	5.66	13.41	<.01	2.42
MR	4.0	2.1	10.5	3.4	6.54	13.32	<.01	2.31
FW	4.9	1.9	10.1	2.8	5.16	11.33	<.01	2.16
DS	3.4	2.5	10.5	3.3	7.02	11.72	<.01	2.40
PS	4.8	2.6	10.8	2.8	5.95	11.82	<.01	2.20
CD	4.7	3.2	10.3	3.0	5.51	9.70	<.01	1.78
SS	4.3	3.1	10.8	3.1	6.48	10.85	<.01	2.09
VCI	64.0	10.8	96.3	15.0	32.27	13.96	<.01	2.47
VSI	67.5	11.0	97.5	12.7	30.06	14.22	<.01	2.53
FRI	68.8	9.7	102.1	14.5	33.32	14.40	<.01	2.70
WMI	66.8	11.7	103.5	15.4	36.68	13.24	<.01	2.68
PSI	68.0	17.5	102.8	15.3	34.79	10.97	<.01	2.12
FSIQ	60.1	10.3	99.6	14.3	39.51	15.42	<.01	3.17
NVI	63.7	9.7	101.1	13.7	37.40	15.95	<.01	3.15
GAI	64.2	8.2	99.2	14.3	35.03	16.20	<.01	3.01
CPI	61.7	15.5	103.1	14.9	41.47	13.39	<.01	2.73

[a] The Standard Difference is the difference of the two test means divided by the square root of the pooled variance, computed using Cohen's (1996) Formula 10.4.

WISC–V abbreviations are: SI = Similarities, VC = Vocabulary, BD = Block Design, VP = Visual Puzzles, MR = Matrix Reasoning, FW = Figure Weights, DS = Digit Span, PS = Picture Span, CD = Coding, SS = Symbol Search, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, NVI = Nonverbal Index, GAI = General Ability Index, CPI = Cognitive Proficiency Index.

The results of this study are consistent with results of prior research involving the paper format of Coding and Symbol Search (Wechsler, 2003, 2014). All mean primary subtest scaled scores and composite scores are significantly lower than those of the matched control group, and all effect sizes are large. The largest effect sizes among primary subtests occur on Visual Puzzles, Digit Span, Vocabulary, and Matrix Reasoning, and the smallest occur on Coding, Block Design, and Symbol Search.

Consistent with prior research on this population that involved the WISC–V in paper format (Wechsler, 2014) and in digital format (Raiford et al., 2014), the variability in performance on the primary subtests and index scores is generally smaller than in the matched control group for all domains. Coding, Symbol Search, and the PSI tend to have larger standard deviations relative to the other subtests and composite scores.

Additional analysis indicates that 93% of the children in the intellectual disability group have FSIQ scores of 75 points or lower versus only 9% of children in the matched control group; 88% have NVI scores of 75 points or lower versus only 5% of children in the matched control group.

The consistency of the present findings with those obtained from a study of the WISC–V administered in paper format (Wechsler, 2014) indicates that Coding, Symbol Search, and the PSI are measuring similar constructs in both formats. Taken together, these results provide strong evidence that when Coding and Symbol Search (and all of the remaining WISC–V primary subtests) are administered in digital format, the WISC–V produces scores that are useful in the assessment of intellectual disability and consistent with results obtained when administering the test in paper format.

Children With Specific Learning Disorders

Specific Learning Disorder-Reading

There is a large body of research evaluating the general and specific cognitive difficulties associated with reading disorder. Although a comprehensive review is beyond the scope of this report, some pertinent findings are highlighted here. In the verbal comprehension domain, studies show that vocabulary knowledge is associated with the development of reading skills (Ouellette, 2006). Children with specific learning disorder-reading (SLD-R) have difficulties with semantic search and retrieval (Booth, Bebko, Burman, & Bitan, 2007), and SLD-R is associated with lower performance on expressive, but not receptive, language measures (Cutting, Materek, Cole, Levine, & Mahone, 2009). Children with reading comprehension deficits show impairments in language functioning compared to controls and children with decoding-only deficits (Catts, Adlof, & Weismer, 2006). In a large sample of children diagnosed with ADHD and LD, verbal comprehension and working memory were the best WISC–III/WISC–IV predictors of reading ability; however, working memory and processing speed scores were the best predictors of a learning disorder (Mayes & Calhoun, 2007). Children diagnosed with SLD-R show reduced verbal working memory (Kibby & Cohen, 2008; Wechsler, 2014). While significant processing speed deficits are present in some studies of individuals with SLD-R (Shanahan et al., 2006), it is typically one of the smallest effect sizes in matched control studies (Wechsler, 2003, 2008, 2014).

The WISC–V in digital format was administered to a sample of children with specific learning disorder-reading. The demographic characteristics of this sample appear in Table 10. Table 13 presents the mean subtest and composite scores for the Specific Learning Disorder-Reading (SLD-R) and matched control groups.

Table 13. Specific Learning Disorder-Reading Compared to Matched Controls

Subtest/ Composite Score	SLD-R		Matched Control		Group Mean Comparison			
	Mean	*SD*	Mean	*SD*	Difference	*t* value	*p* value	Standard Difference[a]
SI	8.9	2.1	10.6	2.4	1.71	2.24	<.05	.76
VC	8.8	3.4	10.3	2.7	1.50	1.76	NS	.49
BD	10.3	2.3	11.0	2.3	.79	1.18	NS	.34
VP	10.1	2.4	10.3	3.3	.18	.20	NS	.06
MR	9.1	2.4	10.9	2.4	1.83	2.43	<.05	.76
FW	9.7	3.0	11.5	2.7	1.82	2.31	<.05	.64
DS	7.7	2.6	11.7	3.0	3.96	5.25	<.01	1.41
PS	8.6	2.3	12.0	2.5	3.33	5.31	<.01	1.39
CD	8.3	2.7	10.0	2.3	1.75	3.06	<.01	.70
SS	9.3	2.6	9.8	2.6	.54	.61	NS	.21
VCI	93.7	13.2	102.4	12.1	8.75	2.30	<.05	.69
VSI	100.2	11.4	104.1	14.0	3.86	.95	NS	.30
FRI	96.0	12.2	107.7	9.9	11.73	3.59	<.01	1.06
WMI	89.8	11.0	111.0	12.8	21.17	6.58	<.01	1.77
PSI	93.3	14.0	99.6	11.7	6.29	1.69	NS	.49
FSIQ	91.5	10.8	106.3	9.1	14.81	5.10	<.01	1.48
NVI	95.7	9.8	107.1	10.9	11.43	3.52	<.01	1.10
GAI	94.9	10.7	105.9	9.4	11.00	3.57	<.01	1.09
CPI	89.7	12.5	106.6	12.3	16.83	4.74	<.01	1.36

[a] The Standard Difference is the difference of the two test means divided by the square root of the pooled variance, computed using Cohen's (1996) Formula 10.4.

WISC–V abbreviations are: SI = Similarities, VC = Vocabulary, BD = Block Design, VP = Visual Puzzles, MR = Matrix Reasoning, FW = Figure Weights, DS = Digit Span, PS = Picture Span, CD = Coding, SS = Symbol Search, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, NVI = Nonverbal Index, GAI = General Ability Index, CPI = Cognitive Proficiency Index.

When compared to a matched control group, children with SLD-R obtained significantly lower mean scores for most primary index scores and the FSIQ. The FSIQ mean difference produced a large effect size, as did those of the FRI and WMI. The remainder of the primary index scores showed small to moderate effect sizes. The largest effect size is observed for the WMI, which is consistent with contemporary research that indicates a relationship between reading achievement and difficulties with multiple components of working memory (Wang & Gathercole, 2013). The mean differences of all global composites (FSIQ, NVI, GAI, and CPI) have large effects. At the subtest level, the largest effect sizes are noted for Digit Span and Picture Span. Consistent with prior results, the mean differences for Coding and Symbol Search produce some of the smallest effect sizes.

The consistency of the present findings with those obtained from a study of the WISC–V administered in paper format (Wechsler, 2014) indicates that Coding, Symbol Search, and the PSI are measuring similar constructs in both formats. These results provide evidence that when Coding and Symbol Search (and all of the remaining WISC–V primary subtests) are administered in digital format, the WISC–V produces scores that are consistent with results obtained from administration of the paper format.

Specific Learning Disorder-Mathematics

Although the research base evaluating math disorder is less extensive than for reading disorder, there is evidence for common cognitive difficulties between the two specific learning disorders, including difficulties in verbal comprehension, working memory, and processing speed (Willcutt et al., 2013). Geary (2011a) found that general cognitive functioning, processing speed, and components of working memory were longitudinal predictors of math achievement. Additionally, early number skills and conceptual reasoning skills predict math achievement (Fuchs, Geary, Compton, Fuchs, Hamlett, & Bryant, 2010); and language, nonverbal reasoning, and attention are significantly related to performance on math word problems (Fuchs, Geary, Compton, Fuchs, Hamlett, Seethaler, et al., 2010; Tolar et al., 2012). Although general cognitive functioning is a predictor for math achievement in typically developing children, it is not a primary cause of math disorder (Geary, 2011b). Difficulties with working memory (Geary, 2010), attention (Raghubar et al., 2009), and semantic-retrieval and visuospatial skills (Cirino, Morris, & Morris, 2007) are related to mathematics difficulties.

The WISC–V in digital format was administered to a sample of children with specific learning disorder-mathematics. The demographic characteristics of this sample appear in Table 10. Table 14 presents the mean subtest and composite scores for the Specific Learning Disorder-Mathematics (SLD-M) and matched control groups.

Table 14. Specific Learning Disorder-Mathematics Compared to Matched Controls

	SLD-M		Matched Control		Group Mean Comparison			
Subtest/ Composite Score	Mean	*SD*	Mean	*SD*	Difference	*t* value	*p* value	Standard Difference[a]
SI	7.8	2.3	10.6	2.2	2.81	4.08	<.01	1.25
VC	8.0	2.3	10.1	2.0	2.14	3.72	<.01	.99
BD	8.6	2.6	10.8	2.1	2.18	2.62	<.05	.92
VP	7.5	2.2	11.0	3.0	3.52	3.98	<.01	1.34
MR	7.9	2.4	11.0	2.5	3.10	4.51	<.01	1.27
FW	8.0	2.0	10.3	2.4	2.33	4.09	<.01	1.05
DS	8.0	2.6	10.8	2.6	2.77	3.18	<.01	1.07
PS	8.5	1.9	10.7	2.6	2.19	3.25	<.01	.96
CD	7.3	3.6	10.1	2.7	2.81	2.93	<.01	.88
SS	7.8	3.2	10.0	2.6	2.18	2.41	<.05	.75
VCI	89.0	9.8	102.1	9.3	13.19	4.70	<.01	1.38
VSI	89.5	11.4	104.8	11.9	15.29	3.54	<.01	1.31
FRI	88.0	10.8	104.4	12.3	16.45	5.27	<.01	1.42
WMI	90.1	10.0	104.1	11.8	14.00	4.19	<.01	1.28
PSI	86.8	18.0	100.4	11.5	13.62	2.85	<.01	.90
FSIQ	85.3	10.1	103.2	9.2	17.89	6.12	<.01	1.85
NVI	85.7	10.7	105.4	10.4	19.72	5.56	<.01	1.87
GAI	86.8	9.1	103.9	10.3	17.11	7.60	<.01	1.76
CPI	85.9	13.3	102.4	10.6	16.50	3.73	<.01	1.37

[a] The Standard Difference is the difference of the two test means divided by the square root of the pooled variance, computed using Cohen's (1996) Formula 10.4.

WISC–V abbreviations are: SI = Similarities, VC = Vocabulary, BD = Block Design, VP = Visual Puzzles, MR = Matrix Reasoning, FW = Figure Weights, DS = Digit Span, PS = Picture Span, CD = Coding, SS = Symbol Search, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, NVI = Nonverbal Index, GAI = General Ability Index, CPI = Cognitive Proficiency Index.

The mean scores for the SLD-M group are significantly lower than the mean scores for the matched control group for all primary index scores and the FSIQ. All effect sizes are large, with the largest observed on the FRI, VSI, and VCI. The smallest effect size among the primary index scores is observed on the PSI, which is consistent with prior results (Wechsler, 2003, 2008, 2014). The mean performance on all subtests is significantly lower in the SLD-M group. The largest effect sizes are observed on Visual Puzzles and Matrix Reasoning.

The consistency of the present findings with those obtained from a study of the WISC–V administered in paper format (Wechsler, 2014) indicates that Coding, Symbol Search, and the PSI are measuring similar constructs in both formats. These results provide evidence that when Coding and Symbol Search (and all of the remaining WISC–V primary subtests) are administered in digital format, the WISC–V produces scores that are consistent with results obtained from administration of the paper format.

Children With Attention-Deficit/Hyperactivity Disorder

Traditional IQ scores generally have not been found useful in discriminating children or adults with ADHD from a nonclinical population; however, their global intellectual functioning displays mild impairment in some studies (Hale et al., 2012; Raiford, Drozdick, & Zhang, 2015; Wechsler, 2014). Children with ADHD show relatively preserved verbal comprehension scores, with lower performance on working memory and processing speed tasks (Hale et al., 2012; Mayes, Calhoun, Chase, Mink, & Stagg, 2009; Mayes, Calhoun, Mayes, & Molitoris, 2012; Wakkinen, 2008; Wechsler, 2012; Zieman, 2010). Fluid Reasoning performance is also lower than matched controls in a number of studies (Raiford et al., 2015; Wechsler, 2012, 2014).

Performance on processing speed tasks is often lower in children with ADHD (Jacobson et al., 2011; Metin et al., 2013; Wechsler, 2003, 2012, 2014). Several studies have found slower response times in ADHD groups versus groups of children without clinical conditions (Chiang, Huang, Gau, & Shang, 2013; Crosbie et al., 2013; Lipszyc & Schachar, 2010; Rosch, Dirlikov, & Mostofsky, 2013).

The WISC–V in digital format was administered to a sample of children with ADHD. The demographic characteristics of this sample appear in Table 10. Table 15 presents the mean subtest and composite scores for the ADHD and matched control groups.

Table 15. Attention-Deficit/Hyperactivity Disorder Compared to Matched Controls

Subtest/ Composite Score	ADHD		Matched Control		Group Mean Comparison			
	Mean	*SD*	Mean	*SD*	Difference	*t* value	*p* value	Standard Difference[a]
SI	10.3	2.6	10.0	2.8	-.29	-.30	NS	-.11
VC	10.0	2.3	10.4	2.9	.35	.51	NS	.13
BD	9.3	2.6	10.3	3.4	1.00	1.06	NS	.33
VP	9.1	1.9	10.4	3.0	1.30	1.67	NS	.52
MR	8.7	2.3	10.0	2.7	1.24	1.94	NS	.49
FW	9.2	3.0	11.5	2.0	2.30	2.80	<.05	.90
DS	9.3	1.1	11.1	3.0	1.76	2.46	<.05	.78
PS	10.0	2.8	11.5	2.7	1.52	1.64	NS	.55
CD	7.9	3.3	10.9	2.7	2.95	4.47	<.01	.98
SS	7.8	3.5	10.0	2.5	2.14	3.00	<.01	.70
VCI	101.7	11.0	101.2	14.7	-.50	-.12	NS	-.04
VSI	94.9	9.1	101.6	16.9	6.65	1.48	NS	.49
FRI	93.2	12.5	103.9	12.2	10.70	2.70	<.05	.87
WMI	97.9	10.1	107.3	12.8	9.38	2.50	<.05	.81
PSI	88.0	17.7	102.0	11.9	14.00	3.89	<.01	.93
FSIQ	94.3	6.4	103.6	14.1	9.32	2.69	<.05	.85
NVI	91.5	8.5	105.1	14.2	13.60	3.55	<.01	1.16
GAI	97.1	8.7	102.1	15.3	4.95	1.18	NS	.40
CPI	91.2	12.0	106.0	10.9	14.76	4.65	<.01	1.29

[a] The Standard Difference is the difference of the two test means divided by the square root of the pooled variance, computed using Cohen's (1996) Formula 10.4.

WISC–V abbreviations are: SI = Similarities, VC = Vocabulary, BD = Block Design, VP = Visual Puzzles, MR = Matrix Reasoning, FW = Figure Weights, DS = Digit Span, PS = Picture Span, CD = Coding, SS = Symbol Search, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, NVI = Nonverbal Index, GAI = General Ability Index, CPI = Cognitive Proficiency Index.

This sample of children with ADHD tested on the WISC–V in digital format had a higher mean age, generally lower parent education level, and a greater proportion of children who are Hispanic than the analogous group that was tested on the WISC–V paper format (Wechsler, 2014). Demographic variations also were observed between the present study and a prior study conducted with the WISC–V in digital format (Raiford et al., 2015). Not surprisingly, these three samples each produced slightly different results; however, the direction of the differences are the same, and the subtest- and composite-level means and effect sizes are generally similar.

The same slight differences are observed when comparing results across the matched-control (nonclinical) groups drawn from the three aforementioned studies, yet the equivalence of the paper and digital versions in nonclinical samples has been established previously (Daniel, 2012, Daniel, Wahlstrom, & Zhang, 2014), and the digital format subtests used for the Raiford et al. (2015) study were identical to those in the present study. Taken together, this pattern of results implies that differences in demographic characteristics (e.g., parent education) or symptom severity are likely responsible for the slight differences across the WISC–V digital and paper ADHD group studies. Comparisons of the WISC–IV paper format ADHD study results (Wechsler, 2003) with analogous published studies on the WISC–IV using groups of children with different demographic characteristics (e.g., Mayes et al., 2009; Zieman, 2010) similarly show sample-related fluctuations.

Significant differences are found between the ADHD and matched control groups on the FRI, WMI, and PSI. The PSI is the lowest of all mean primary index scores for the ADHD group. These results are remarkably similar to those of the paper version. Both Coding and Symbol Search show statistically significant differences between the ADHD and matched control groups, as do Figure Weights and Digit Span. The largest effects are observed for Coding and Figure Weights. Significant differences are also present on Symbol Search and Digit Span. These results are indicative of the anticipated difficulties in working memory and processing speed.

The consistency of the present findings with those obtained from a study of the WISC–V administered in paper format (Wechsler, 2014) indicates that Coding, Symbol Search, and the PSI are measuring similar constructs in both formats. These results provide evidence that when Coding and Symbol Search (and all of the remaining WISC–V primary subtests) are administered in digital format, the WISC–V produces scores that are consistent with results obtained from administration of the paper format.

Children With Autism Spectrum Disorder With Accompanying Language Impairment

Previous investigations suggest that the general intellectual functioning of individuals with autism spectrum disorder is lower than that of matched controls; however, these studies suggest a pattern of strengths and weaknesses. Multiple studies demonstrate lower scores on measures of general intellectual functioning but relatively better performance on measures of fluid reasoning (Dawson, Soulières, Gernsbacher, & Mottron, 2007; Mayes & Calhoun, 2008; Stevenson, 2011). Performance on verbal tasks is typically lower for most children with autism spectrum disorder with accompanying language impairment (ASD-L) than typically developing children (Joseph, Tager-Flusberg, & Lord, 2002; Klinger, O'Kelley, Mussey, Goldstein, & DeVries, 2012; Mayes & Calhoun, 2008; Raiford et al., 2015; Wechsler, 2003, 2012, 2014). A typical pattern of performance on the Verbal Comprehension subtests has emerged: the highest score is obtained on Similarities, which involves fluid reasoning, and the lowest score on Comprehension, which requires some social judgment—a weakness in individuals with autism spectrum disorder (Mayes & Calhoun, 2008; Zayat, Kalb, & Wodka, 2011). In addition, some studies show relative strengths on visual spatial tasks for children with autistic disorder (Mayes & Calhoun, 2008; Raiford et al., 2015; Wechsler, 2003, 2012, 2014). A prior study indicated that a group of children with autism spectrum disorder with language impairment showed very similar performance on the WISC–V in digital format (Raiford et al., 2015) relative to a group of children with the same condition tested with the paper format (Wechsler, 2014).

The WISC–V in digital format was administered to a sample of children with autism spectrum disorder with accompanying language impairment. The demographic characteristics of this sample appear in Table 10. Table 16 presents the mean subtest and composite scores for the Autism Spectrum Disorder With Accompanying Language Impairment (ASD-L) and matched control groups.

Table 16. Autism Spectrum Disorder With Language Impairment Compared to Matched Controls

Subtest/ Composite Score	ASD-L Mean	ASD-L *SD*	Matched Control Mean	Matched Control *SD*	Group Mean Comparison Difference	*t* value	*p* value	Standard Difference[a]
SI	7.6	3.5	10.0	1.9	2.38	3.36	<.01	.85
VC	7.2	3.8	10.2	2.4	3.04	3.68	<.01	.96
BD	9.0	3.7	10.3	3.4	1.23	1.34	NS	.35
VP	9.3	3.6	10.2	2.6	.96	1.09	NS	.31
MR	8.7	3.9	11.1	2.4	2.46	3.06	<.01	.76
FW	8.2	4.1	11.6	2.2	3.42	3.97	<.01	1.04
DS	6.5	4.1	11.3	3.3	4.77	4.21	<.01	1.28
PS	8.3	3.2	11.4	2.6	3.16	3.70	<.01	1.08
CD	6.5	3.6	10.2	2.8	3.64	4.66	<.01	1.13
SS	7.4	3.9	10.6	3.1	3.24	3.52	<.01	.92
VCI	86.0	19.1	100.4	9.9	14.40	3.89	<.01	.95
VSI	95.5	19.6	100.9	14.6	5.44	1.13	NS	.31
FRI	91.0	19.6	107.5	11.4	16.54	4.53	<.01	1.03
WMI	85.6	19.2	107.7	14.2	22.16	4.27	<.01	1.31
PSI	82.0	20.7	102.6	14.7	20.63	4.64	<.01	1.15
FSIQ	81.4	18.4	105.0	10.3	23.61	7.48	<.01	1.58
NVI	86.3	18.8	106.3	11.9	19.96	5.55	<.01	1.27
GAI	87.5	18.0	103.8	10.1	16.33	4.93	<.01	1.12
CPI	80.3	22.1	107.0	13.9	26.79	5.78	<.01	1.45

[a] The Standard Difference is the difference of the two test means divided by the square root of the pooled variance, computed using Cohen's (1996) Formula 10.4.

WISC–V abbreviations are: SI = Similarities, VC = Vocabulary, BD = Block Design, VP = Visual Puzzles, MR = Matrix Reasoning, FW = Figure Weights, DS = Digit Span, PS = Picture Span, CD = Coding, SS = Symbol Search, VCI = Verbal Comprehension Index, VSI = Visual Spatial Index, FRI = Fluid Reasoning Index, WMI = Working Memory Index, PSI = Processing Speed Index, FSIQ = Full Scale IQ, NVI = Nonverbal Index, GAI = General Ability Index, CPI = Cognitive Proficiency Index.

For the ASD-L group, all mean primary index scores except the VSI are significantly lower than the corresponding means of the matched control group, and those with significant differences also show large effect sizes. A small effect is present on the VSI. Consistent with previous research, the VSI mean is higher and the mean difference shows a smaller effect size relative to the VCI. The WMI produces the largest effect size of the primary index scores, followed by the PSI. This is consistent with research demonstrating weaknesses in working memory and processing speed in children with autism spectrum disorders (Boucher & Mayes, 2012; Corbett, Constantine, Hendren, Rocke, & Ozonoff, 2009; Englund, Decker, Allen, & Roberts, 2014; Mayes & Calhoun, 2007; Raiford et al., 2015; Wechsler, 2003, 2012, 2014).

At the subtest level, with the exception of Block Design and Visual Puzzles, all mean scaled scores are significantly lower in the ASD-L group compared with the matched control group. The largest effect sizes occur on Digit Span, Coding, Picture Span, and Figure Weights.

These results replicate previous research indicating global cognitive deficits, relatively weaker verbal task performance, and relatively higher performance on visual spatial tasks (Barbeau, Soulières, Dawson, Zeffiro, & Mottron, 2013; Klinger et al., 2012; Mayes & Calhoun, 2008; Soulières, Dawson, Gernsbacher, & Mottron, 2011; Raiford et al., 2015; Wechsler, 2003, 2012, 2014). The consistency of the present findings with those obtained from a study of the WISC–V administered in paper format (Wechsler, 2014) indicates that Coding, Symbol Search, and the PSI are measuring similar constructs in both formats. These results provide evidence that when Coding and Symbol Search (and all of the remaining WISC–V primary subtests) are administered in digital format, the WISC–V produces scores that are consistent with results obtained from administration of the paper format.

Children With Motor Impairment

Prior research with children with motor impairment indicated significant mean differences on Coding, Symbol Search, and the PSI relative to matched controls (Wechsler, 2003). All of these effect sizes were large. The mean differences on Similarities, Vocabulary, and the VCI were not significant, but the mean difference for the VCI had a small effect size.

Coding and Symbol Search, along with the two primary Verbal Comprehension subtests (i.e., Similarities and Vocabulary) were administered in digital format to a sample of children with significant motor impairment. Children with significant motor impairment due to cerebral palsy were included in the study if no concurrent diagnosis of intellectual disability was present. The primary purpose of this study was to illustrate typical performance on Coding among children with motor impairment with touch responses rather than written responses. The demographic characteristics of this sample appear in Table 10. Table 17 presents the mean subtest and composite scores for the Motor Impairment and matched control groups.

Table 17. Motor Impairment Compared to Matched Controls

Subtest/ Composite Score	Motor Impairment		Matched Control		Group Mean Comparison			
	Mean	*SD*	Mean	*SD*	Difference	*t* value	*p* value	Standard Difference[a]
SI	9.3	3.7	10.8	3.9	1.50	1.23	NS	.39
VC	10.2	2.3	11.1	2.4	0.93	1.33	NS	.40
CD	7.1	4.0	10.2	4.0	3.07	2.21	<.05	.77
SS	7.1	3.7	10.1	3.4	3.07	3.03	<.01	.86
VCI	98.7	14.6	105.5	16.6	6.79	1.38	NS	.43
PSI	83.1	20.4	101.1	18.2	17.93	2.88	<.05	.93

[a] The Standard Difference is the difference of the two test means divided by the square root of the pooled variance, computed using Cohen's (1996) Formula 10.4.

WISC–V abbreviations are: SI = Similarities, VC = Vocabulary, CD = Coding, SS = Symbol Search, VCI = Verbal Comprehension Index, PSI = Processing Speed Index.

For the Motor Impairment group, all results are remarkably similar to those of the WISC–IV paper format study (Wechsler, 2003) despite the reduced graphomotor skill requirements for Coding in the digital format. Coding, Symbol Search, and the PSI are significantly lower than the corresponding means of the matched control group. The Coding and Symbol Search means are the lowest of these subtest scores, consistent with prior results obtained with the WISC–IV paper format (Wechsler, 2003). The Coding mean produces a moderate effect size, and the Symbol Search and PSI mean differences produce large effect sizes. Consistent with previous research, the Similarities, Vocabulary, and VCI mean differences are not significant. The

Similarities, Vocabulary, and VCI mean differences show small effect sizes. The present findings suggest that administering Coding, Symbol Search, Similarities, and Vocabulary in digital format to children with motor impairment produces similar results as the paper format at the subtest and index score level.

Summary of Special Group Studies

Results from the special group studies provide strong support for the validity and clinical utility of Coding and Symbol Search in digital format, and for the WISC–V digital format in general. Results are consistent with previous research and theoretical foundations. Future independent investigations utilizing the WISC–V in different clinical settings and populations will provide additional evidence of the scale's utility in digital format when used as part of a comprehensive clinical evaluation for diagnosis and intervention purposes.

Interpretation

A slight adjustment to interpretive statements may be helpful when discussing score differences between Coding and Symbol Search and composite scores that involve Coding. With the digital format of Coding, the most salient point is that some of the graphomotor demands of Coding have been removed; therefore, the usual reference or hypothesis that differences between Coding and Symbol Search may be attributable to the graphomotor demands is likely not warranted. However, psychomotor speed continues to be involved with both subtests. Given the continued low performance of the motor impairment group, it is possible that observed differences between Coding and Symbol Search may be more related to task complexity and associative learning as opposed to graphomotor speed.

In addition, Coding responses are now collected within a multiple-choice format, so rotation errors are no longer possible. Therefore, base rates for rotation errors cannot be provided for the digital format. To account for these changes, adjustments have already been made to the interpretive reports that can be generated within Q-interactive.

Cancellation, which is still administered using a paper response booklet, still may be substituted for Coding to obtain the FSIQ. As with any substitution, it is important to note the impact on interpretation of the FSIQ. Specifically, Cancellation has graphomotor demands whereas Coding does not any longer. However, both subtests have been show to load on Processing Speed, and substitution continues to remain an appropriate use for the Cancellation subtest.

References

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: Author.

Barbeau, E. B., Soulières, I., Dawson, M., Zeffiro, T. A., & Mottron, L. (2013). The level and nature of autistic intelligence III: Inspection time. *Journal of Abnormal Psychology, 122*(1), 295–301. doi:10.1037/a0029984

Booth, J. R., Bebko, G., Burman, D. D., & Bitan, T. (2007). Children with reading disorder show modality independent brain abnormalities during semantic tasks. *Neuropsychologia, 45*(4), 775–783. doi:10.1016/j.neuropsychologia.2006.08.015

Boucher, J., & Mayes, A. (2012). Memory in ASD: Have we been barking up the wrong tree? *Autism, 16*(6), 603–611. doi:10.1177/1362361311417738

Calero, M. D., García-Martín, M. B., Jiménez, M. I., Kazén, M., & Araque, A. (2007). Self-regulation advantage for high-IQ children: Findings from a research study. *Learning and Individual Differences, 17*, 328–343.

Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: A case for the simple view of reading. *Journal of Speech, Language, and Hearing Research, 49*(2), 278–293. doi:10.1044/1092-4388(2006/023)

Chiang, H.-L., Huang, L.-W., Gau, S. S.-F., & Shang, C.-Y. (2013). Associations of symptoms and subtypes of attention-deficit hyperactivity disorder with visuospatial planning ability in youth. *Research in Developmental Disabilities, 34*(9), 2986–2995. doi:10.1016/j.ridd.2013.06.020

Cirino, P. T., Morris, M. K., & Morris, R. D. (2007). Semantic, executive, and visuospatial abilities in mathematical reasoning of referred college students. *Assessment, 14*(1), 94–104. doi:10.1177/1073191106291487

Cohen, B. H. (1996). *Explaining psychological statistics.* Pacific Grove, CA: Brooks & Cole.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Corbett, B. A., Constantine, L. J., Hendren, R., Rocke, D., & Ozonoff, S. (2009). Examining executive functioning in children with autism spectrum disorder, attention deficit hyperactivity disorder and typical development. *Psychiatry Research, 166*(2–3), 210–222. doi:10.1016/j.psychres.2008.02.005

Crosbie, J., Arnold, P., Paterson, A., Swanson, J., Dupuis, A., Li, X.,…Schachar, R. J. (2013). Response inhibition and ADHD traits: Correlates and heritability in a community sample. *Journal of Abnormal Child Psychology, 41*, 497–507. doi:10.1007/s10802-012-9693-9

Cutting, L. E., Materek, A., Cole, C. A. S., Levine, T. M., & Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of Dyslexia, 59*(1), 34–54. doi:10.1007/s11881-009-0022-0

Daniel, M. H. (2012). *Equivalence of Q-interactive administered cognitive tasks: WISC–IV* (Q-interactive Technical Report 2). Bloomington, MN: Pearson. Retrieved from http://www.helloq.com/content/dam/ped/ani/us/helloq/media/Technical%20Report%202_WISC-IV_Final.pdf

Daniel, M. H., Wahlstrom, D., & Zhang, O. (2014). *Equivalence of Q-interactive and paper administration of cognitive tasks: WISC®–V* (Q-interactive Technical Report 7). Bloomington, MN: Pearson.

Dawson, M., Soulieres, I., Gernsbacher, M. A., & Mottron, L. (2007). The level and nature of autistic intelligence. *Psychological Science, 18*(8), 657–662. doi:10.1111/j.1467-9280.2007.01954.x

Dorans, N. J. (2004). Equating, concordance, and expectation. *Applied psychological measurement, 28*(4), 227–246.

Englund, J. A., Decker, S. L., Allen, R. A., & Roberts, A. M. (2014). Common cognitive deficits in children with attention-deficit/hyperactivity disorder and autism: Working memory and visual-motor integration. *Journal of Psychoeducational Assessment, 32*(2), 95–106. doi:10.1177/0734282913505074

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. D. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development, 81*(5), 1520–1533. doi:10.1111/j.1467-8624.2010.01489.x

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., . . . Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities? *Developmental Psychology, 46*(6), 1731–1746. doi:10.1037/a0020662

Geary, D. C. (2010). Mathematical disabilities: Reflections on cognitive, neuropsychological, and genetic components. *Learning and Individual Differences, 20*(2), 130. doi:10.1016/j.lindif.2009.10.008

Geary, D. C. (2011a). Cognitive predictors of achievement growth in mathematics: A 5-year longitudinal study. *Developmental Psychology, 47*(6), 1539–1552. doi:10.1037/a0025510.

Geary, D. C. (2011b). Consequences, characteristics, and causes of mathematical learning disabilities and persistent low achievement in mathematics. *Journal of Developmental and Behavioral Pediatrics, 32*(3), 250–263. doi:10.1097/DBP.0b013e318209edef.

Gordon, S., Duff, S., Davidson, T., & Whitaker, S. (2010). Comparison of the WAIS–III and WISC–IV in 16-year-old special education students. *Journal of Applied Research in Intellectual Disabilities, 23*, 197–200. doi:10.1111/j.1468-3148-2009-00538.x

Hale, J. B., Yim, M., Schneider, A. N., Wilcox, G., Henzel, J. N., & Dixon, S. G. (2012). Cognitive and neuropsychological assessment of attention-deficit/hyperactivity disorder: Redefining a disruptive behavior disorder. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 687–707). New York, NY: Guilford Press.

Jacobson, L. A., Ryan, M., Martin, R. B., Ewen, J., Mostofsky, S. H., Denckla, M. B., & Mahone, E. M. (2011). Working memory influences processing speed and reading fluency in ADHD. *Child Neuropsychology, 17*(3), 209–224. doi:10.1080/09297049.2010.532204

Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, 43*(6), 807–821. doi:10.1111/1469-7610.00092

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman assessment battery for children* (2nd ed.). Bloomington, MN: NCS Pearson.

Kibby, M. Y., & Cohen, M. J. (2008). Memory functioning in children with reading disabilities and/or attention deficit/hyperactivity disorder: A clinical investigation of their working memory and long-term memory functioning. *Child Neuropsychology, 14*(6), 525–546. doi:10.1080/09297040701821752

Klinger, L. G., O'Kelley, S. E., Mussey, J. L., Goldstein, S., & DeVries, M. (2012). Assessment of intellectual functioning in autism spectrum disorder. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed., pp. 670–686). New York, NY: Guilford Press.

Lipszyc, J., & Schachar, R. (2010). Inhibitory control and psychopathology: A meta-analysis of studies using the stop signal task. *Journal of the International Neuropsychological Society, 16*(6), 1064–1076. doi:10.1017/S1355617710000895

Mayes, S. D., & Calhoun, S. L. (2007). Learning, attention, writing, and processing speed in typical children and children with ADHD, autism, anxiety, depression, and oppositional-defiant disorder. *Child Neuropsychology, 13*, 469–493. doi:10.1080/09297040601112773

Mayes, S. D., & Calhoun, S. L. (2008). WISC–IV and WIAT–II profiles in children with high-functioning autism. *Journal of Autism and Developmental Disorders, 38*, 429–439. doi:10.1007/s10803-007- 0410-4

Mayes, S. D., Calhoun, S. L., Chase, G. A., Mink, D. M., & Stagg, R. E. (2009). ADHD subtypes and co-occurring anxiety, depression, and oppositional-defiant disorder: Differences in Gordon diagnostic system and Wechsler working memory and processing speed index scores. *Journal of Attention Disorders, 12*(6), 540–550. doi:10.1177/1087054708320402

Mayes, S. D., Calhoun, S. L., Mayes, R. D., & Molitoris, S. (2012). Autism and ADHD: Overlapping and discriminating symptoms. *Research in Autism Spectrum Disorders, 6*, 277–285. doi:10.1016/j.rasd.2011.05.009

Metin, B., Roeyers, H., Wiersema, J. R., van der Meere, J. J., Thompson, M., & Sonuga-Barke, E. (2013). ADHD performance reflects inefficient but not impulsive information processing: A diffusion model analysis. *Neuropsychology, 27*(2), 193–200.

Nunes, M. M., Honjo, R. S., Dutra, R. L., Amaral, V. A. S., Oh, H. K., Bertola, D. R., … Teixeira, M. C. T. V. (2012). Assessment of intellectual and visuo-spatial abilities in children and adults with Williams syndrome. *Universitas Psychologica, 12*(2), 581–589.

Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554 –566. doi:10.1037/0022-0663.98.3.554

Periathiruvadi, S., & Rinn, A. N. (2012). Technology in gifted education: A review of best practices and empirical research. *Journal of Research on Technology in Education, 45*(2), 153–169.

Raghubar, K., Cirino, P., Barnes, M., Ewing-Cobbs, L., Fletcher, J., & Fuchs, L. (2009). Errors in multi-digit arithmetic and behavioral inattention in children with math difficulties. *Journal of Learning Disabilities, 42*(4), 356–371. doi:10.1177/0022219409335211

Raiford, S. E., Drozdick, L. W., & Zhang, O. (2015). *Q-interactive special group studies: The WISC–V and children with autism spectrum disorder and accompanying language impairment or attention-deficit/hyperactivity disorder* (Q-interactive Technical Report 11). Bloomington, MN: Pearson. Retrieved from http://images.pearsonclinical.com/images/assets/WISC-V/Q-i-TR11_WISC-V_ADHDAUTL_FNL.pdf

Raiford, S. E., Holdnack, J. A., Drozdick, L. W., & Zhang, O. (2014). *Q-interactive special group studies: The WISC–V and children with intellectual giftedness and intellectual disability* (Q-interactive Technical Report 9). Bloomington, MN: Pearson. Retrieved from http://www.helloq.com/content/dam/ped/ani/us/helloq/media/Technical_Report_9_WISC-V_Children_with_Intellectual_Giftedness_and_Intellectual_Disability.pdf

Raiford, S. E., Weiss, L. G., Rolfhus, E., & Coalson, D. (2005). *General ability index* (WISC–IV Technical Report 4). San Antonio, TX: Pearson. Retrieved from http://www.pearsonassessments.com/NR/rdonlyres/1439CDFE- 6980-435F-93DA-05888C7CC082/0/80720_WISCIV_Hr_r4.pdf

Rimm, S., Gilman, B., & Silverman, L. (2008). Nontraditional applications of traditional testing. In J. L. VanTassel-Baska (Ed.), *Alternative assessments with gifted and talented students* (pp. 175–202). Waco, TX: Prufrock Press.

Rosch, K. S., Dirlikov, B., & Mostofsky, S. H. (2013). Increased intrasubject variability in boys with ADHD across tests of motor and cognitive control. *Journal of Abnormal Child Psychology, 41*(3), 485–495.

Rowe, E. W., Kingsley, J. M., & Thompson, D. F. (2010). Predictive ability of the general ability index (GAI) versus the full scale IQ among gifted referrals. *School Psychology Quarterly, 25*(2), 119–128. doi:10.1037/a0020148

Shanahan, M. A., Pennington, B. F., Yerys, B. E., Scott, A., Boada, R., Willcutt, E. G., . . . DeFries, J. C. (2006). Processing speed deficits in attention deficit/hyperactivity disorder and reading disability. *Journal of Abnormal Child Psychology, 34*, 585–602. doi:10.1007/s10802-006-9037-8

Soulières, I., Dawson, M., Gernsbacher, M. A., & Mottron, L. (2011). The level and nature of autistic intelligence II: What about Asperger syndrome? *PloS ONE, 6*(9), e25372. doi:10.1371/journal. pone.0025372

Steiner, H. H. (2006). A microgenetic analysis of strategic variability in gifted and average-ability children. *The Gifted Child Quarterly, 50*(1), 62–74.

Stevenson, J. L. (2011). *Autistic cognition: Effects of test domain and reasoning level* (Doctoral dissertation). Retrieved from ProQuest Dissertations and Theses. (UMI No. 3486771)

Tolar, T. D., Fuchs, L., Cirino, P. T., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2012). Predicting development of mathematical word problem solving across the intermediate grades. *Journal of Educational Psychology, 104*(4), 1083–1093. doi:10.1037/a0029020

Wakkinen, H. B. (2008). *Maximizing resources to gain information about clients: Profile analysis, configural frequency analysis, and the WISC–IV*. Retrieved from ProQuest Dissertations and Theses. (UMI No. 3322469)

Wang, S., & Gathercole, S. E. (2013). Working memory deficits in children with reading difficulties: Memory span and dual task coordination. *Journal of Experimental Child Psychology, 115*(1), 188–197. doi:10.1016/j.jecp.2012.11.015

Wechsler, D. (2003). *Wechsler intelligence scale for children* (4th ed.). Bloomington, MN: Pearson.

Wechsler, D. (2008). *Wechsler adult intelligence scale* (4th ed.). Bloomington, MN: Pearson.

Wechsler, D. (2012). *Wechsler preschool and primary scale of intelligence* (4th ed.). Bloomington, MN: Pearson.

Wechsler, D. (2014). *Wechsler intelligence scale for children* (5th ed.). Bloomington, MN: Pearson.

Wilkins, C., Rolfhus, E., Weiss, L., & Zhu, J. J. (2005, April). *A new method for calibrating translated tests with small sample sizes*. Paper presented at the 2005 annual meeting of the American Educational Research Association, Montreal, Canada.

Willcutt, E. G., Petrill, S. A., Wu, S., Boada, R., DeFries, J. C., Olson, R. K., & Pennington, B. F. (2013). Comorbidity between reading disability and math disability: Concurrent psychopathology, functional impairment, and neuropsychological functioning. *Journal of Learning Disabilities, 46*(6), 500–516. doi:10.1177/0022219413477476

Zayat, M., Kalb, L., & Wodka, E. L. (2011). Brief report: Performance pattern differences between children with autism spectrum disorders and attention deficit-hyperactivity disorder on measures of verbal intelligence. *Journal of Autism and Developmental Disorders, 41*(12), 1743–1747. doi:10.1007/s10803- 011-1207-z

Zhu, J. J. & Chen, H-Y. (2011). Utility of inferential norming with smaller sample sizes. *Journal of Psychoeducational Assessment, 29*(6), 570–580.

Zieman, S. F. X., Jr. (2010). *Performance analysis on the WISC–IV working memory and processing speed index among ADHD subtypes*. (Doctoral dissertation). Retrieved from WorldCat. (Accession No. 526695555)