Woodcock Johnson IV Tests of Cognitive Abilities (WJ IV COG): Updated Guidance on using the battery within a diagnostic assessment for SpLDs February 2022	
Author:	Shrank F, Mather N, McGrew K, Wendling B, Woodcock R
Publisher:	Riverside Insights 2014
Date of Standardisation:	2009-2012 (US sample)
Age Range:	2-90+
Access:	High: See Riverside Insight's explanation of qualification levels on their website at https://cms.riversideinsights.com/uploads/a07bc02f5b56459bbd1a6978be44925a.pdf Riverside Insights have indicated that a Specialist Teacher Assessor with an APC would be suitably qualified to use the WJIV Tests of Cognitive Abilities.
Sample	Over 7000
Cost	(From Education Elephant website) c€1,785.00 for the complete kit. Subsequently c€94 for additional 25 Student Response Booklets and €214 for 25 Examiner Test Records plus scoring, plus costs of importing

Items which have been updated since the previous (Feb 2021) guidance have been highlighted in yellow.

OVERVIEW

Strengths

- The battery specifically states that it is designed as a tool chest and that assessors should choose individual subtests that are relevant to the assessment. "*An examiner seldom needs to administer all tests or complete all interpretive options for a single person. The importance of selecting testing becomes apparent as the examiner gains familiarity with the WJ IV COG. An analogy to craftsmanship is appropriate: the WJ IV COG provides an extensive tool chest that can be used selectively by a variety of skilled assessment professionals. Different assessments require different combinations of tools."* P3 Examiner Manual

- WJIV could potentially be used to give evaluations of vocabulary knowledge, a range of reasoning skills, processing speed, working memory, phonological awareness and phonological memory. These tests are co-normed – which aids comparison between cognitive scores. Further information about the individual

subtests (what they measure and how they could fit into a diagnostic assessment) is available on pages 5-14 below.

- Test 1 Oral Vocabulary, Test 8 General Information, Test 2 Number Series and Test 9 Concept Formation together produce a cluster score which is called the Gf-Gc Composite. This composite "*is designed to be an estimate of intellectual ability based on the two highest order factors: comprehension knowledge [Crystallized Intelligence/Verbal Ability] and fluid reasoning [Non-Verbal Ability].*" If the Gf-Gc Composite is calculated, the online scoring platform is then able to make complex comparisons with other cognitive abilities. All comparisons are corrected for regression to the mean. The comparson offers information on:
  - The size of any discrepancy (in terms of standard deviations – SD).
  - Whether the discrepancy indicates a strength or weakness
  - A discrepancy percentile rank (PR) which shows the percentage of the population identified as possessing a discrepancy of that magnitude.

	STANDARD SCORES			DISCREPANCY		Interpretation at
COMPARISONS	Actual	Predicted	Difference	PR	SD	+ or - 1.50 SD
Gf-Gc Composite/Other Ability Comparisons						
S-TERM WORK MEM (*Gwm*)	91	112	-21	3	-1.88	Weakness
COG PROCESS SPEED (*Gs*)	87	109	-22	6	-1.52	Weakness
AUDITORY PROCESS (*Ga*)	94	112	-18	6	-1.58	Weakness
COGNITIVE EFFICIENCY	81	111	-30	1	-2.47	Weakness

Example of Gf-GC Composite/Other Ability Comparisons in an individual's score report:

- Whilst Cluster (composite) scores are provided, there is no requirement that a particular core set of subtests should be administered or that all the subtests within a cluster should be administered. Individual subtest scores have high reliability in their own right.

SUBTEST	MEDIAN RELIABILITY	
	5 to 19 age range	Adult age range
Test 1: Oral Vocabulary	.89	.92
Test 2: Number Series	.91	.90
Test 3 Verbal Attention	.86	.83
Test 4 Letter Pattern Matching	.88	.91
Test 5: Phonological Processing	.83	.90
Test 6 Story Recall	.93	.91
Test 7 Visualisation	.83	.87
Test 8: General Information	.84	.91
Test 9 Concept Formation	.93	.95
Test 10 Numbers Reversed	.84	.91

Test 11 Number Pattern Matching	.85/.84	.88
Test 12: Nonword Repetition	.90	.90
Test 13 Visual Auditory Learning	.96	.98
Test 14 Picture Recognition	.71	.73
Test 15: Analysis-Synthesis	.92	.94
Test 16 Object-Number Sequencing	.89	.87
Test 17 Pair Cancellation	.89	.95

- Timing: only tests which measure speed are timed. Other tests may include guidelines on how long to wait before encouraging the individual to respond or move on to the next item.
- All WJIV batteries have been normed on the same sample. So WJIV Cog can be used alongside WJIV Achievement.
- A number of the tests are designed as controlled learning tasks where knowledge is built up and applied cumulatively through the test. This offers the potential to examine capacity to learn, with less interference from acquired knowledge and past learning experiences
- Although it is an American test, the number of Americanisms is relatively few
- It covers the full age range from 2 up to 90 years
- It was standardised in 2009-2012 and published in 2014. Standardisation is robust.
- Instructions are clear and easy to follow.
- There are varying Suggested Starting Points based on estimates of the examinee's present ability (by school grade). However, the assessor can select a starting point that seems most appropriate for the examinee. Basals and Ceilings are provided for most tests, except a few where everyone starts at same point or the item is timed. Usually (basals) are the 4, 5 or 6 lowest numbered items on the page correct or (ceilings) the last 4, 5, or 6 on the highest numbered items on the page incorrect. Clear instructions provided as to basals and ceilings and how to apply them.
- Document available from publisher's website about online administration of this battery. Digital copies of the examinee facing stimuli materials are available on the Resources tab of the Riverside Score platform to all customers who have an account. They were developed so an assessor may more easily screen share the stimuli to the examinee, without showing the assessor side of the pages. The assessor needs to have the actual easel in front of them to have access to the instructions. Riverside Insights provides brief guidance on remote use of the tests at

https://cms.riversideinsights.com/uploads/735e4b05cf5142c695b3b4b5d3fb42a1.pdf .

Issues to consider

- The WJIV Cog provides limited evaluation of reasoning with language (rather than pictures or numbers)
- The WJIV Cog does not provide a measure of Rapid Naming. If a measure of Rapid Naming is required it would need to be taken from another battery. Recent discussions indicate that there is still much uncertainty about what rapid naming measures and why we use it within an assessment. It is widely recognised to be a robust early predictor of reading fluency.
- In some respects the WJIV Cog varies from WRIT and WISC/WAIS. Studies were done comparing the construct validity of WJIV COG against WAIS IV and WISC IV.

WAIS IV Measure	WJ IV COG Measure	Correlation
General Ability Index	Gf-Gc Composite	.75
Verbal Comprehension Index	Comprehension-Knowledge (Gc)	.74
Perceptual Reasoning Index	Fluid Reasoning (Gf)	.57
Working Memory Index	Short-term Working Memory (Gsm)	.67
Processing Speed Index	Processing Speed (Gs)	.44

"*Research indicates the WAIS IV Perceptual Reasoning Index is not as strong a measure of Gf as the corresponding WISC-IV index: the WAIS IV Perceptual Reasoning Index is more a measure of Gv – composed of two Gv tests (Block Design and Visual Puzzles) and one Gf text (Matrix Reasoning).*"

WISC IV Measure	WJ IV COG Measure	Correlation
General Ability Index	Gf-Gc Composite	.83
Verbal Comprehension Index	Comprehension-Knowledge (Gc)	.79
Perceptual Reasoning Index	Fluid Reasoning (Gf)	.70
Working Memory Index	Short-term Working Memory (Gsm)	.72
Processing Speed Index	Processing Speed (Gs)	.55

The manual notes the moderate correlation between Processing Speed measures, and suggests that "*although both clusters measure Gs, the WJ IV and WISC-IV processing speed composites each also measure unique abilities not measured by the other battery's composite score*."

- Cost may seem high but it potentially replaces 4 batteries of tests.
- It is a very large battery of tests. Average time to administer the entire battery would be 90 minutes (though again this should be compared against time taken to administer tests of ability, memory, phonology and processing speed). The manual reports that on average it takes 35 minutes to administer all 7 of Tests 1 – 7 and approximately 5 minutes for each additional test administered.
- Not all the subtests within the battery are core to a diagnostic assessment, and assessors may need additional guidance/training in which subtests to use.
- As with WJIV Achievement, scoring is done entirely online, and there are no published norm tables. This means that the scoring lacks transparency. It considerably limits the assessor's scope to evaluate an examinee's scores according to where they lie within the distribution for a particular subtest. Difficult to get a feel for what the ceiling levels would be for each subtests at each age/educational level. The Technical Manual does include graphs showing typical developmental curves across the whole age range – and these could be used to develop a feel for what is typical at a particular age. Further information about the advantages and disadvantages of online scoring can be found in "**Guidance on the use of digital scoring platforms supplied with certain assessment materials"** on the Downloads page of the SASC website, or at https://sasc.org.uk/SASCDocuments/Guidance%20on%20the%20use%20of%20digital%20assessment%20scoring%20platforms.pdf

WJIV TESTS OF COGNITIVE ABILITIES: HOW SUBTESTS RELATE TO THE DIAGNOSTIC REPORT FORMAT

The following offers suggestions as to which subtests might become core elements within the diagnostic report format, and which ones could be regarded as useful in certain circumstances.

A) Verbal Abilities

Test 1 Oral Vocabulary (p3, p25)

- Synonyms – examinee listens to a word and then provides a single word with the same or similar meaning. 31 items.
- Antonyms - examinee listens to a word and then provides a single word with an opposite meaning. 32 items.

It is necessary to administer both tests to achieve an Oral Vocabulary score

Test 8 General Information.

- Where: "Where would you find . . . (an object)? 22 items. The examinee can answer with a word or a short phrase.
- What: "What would you do with . , . . (an object)? 22 items. The examinee can answer with a word or a short phrase.

It is necessary to administer both tests to achieve an Oral Vocabulary score

These four form the Comprehension – Knowledge cluster (Gc) also known as crystallised intelligence. They explore abilities that have been developed largely through investment of time, talent and resources during education and general life experiences. Gc includes

- Breadth and depth of acquired knowledge (mostly language based)
- Ability to communicate one's knowledge (especially verbally)
- Ability to reason using previously learned experiences or procedures

Strengths	Issues to consider
• Quick and easy to administer. Relatively easy to note responses and evaluate their accuracy. • Measures depth and breadth of acquired vocabulary knowledge, which makes it useful for identifying strengths/weaknesses in oral vocabulary knowledge, and/or identifying disparity between well-developed oral vocabulary knowledge, versus weaker reading and writing skills. • Oral Vocabulary's requirement that the answer be given as a single word makes it a very challenging test which places heavy demands on the	• The subtests make limited demands upon ability to communicate in phrases or sentences or the ability to reason. Where assessors feel an examinee's profile requires further exploration of expressive language skills, it may be necessary to use additional tests from other batteries and/or to use informal qualitative observation of verbal skills throughout the assessment • The subtests involve some reasoning but the type of Inductive reasoning measured in tests such was WISC/WAIT Similarities or

examinee's wider vocabulary knowledge. It will give a strong indication of how likely the examinee is to be able to select vocabulary to express ideas *accurately* and *succinctly.* The vocabulary covered tends to relate to abstract concepts. • General Information tests focus more on being able to demonstrate understanding of the vocabulary, and the vocabulary included tends to relate to concrete objects. • Relatively little US cultural interference. • Items tend to be general rather than subject specific, though there are some more subject specific items towards the end of Where and What.	WRIT Verbal Analogies is covered more extensively in the Fluid Reasoning tests • One item in Oral Vocabulary can have different pronunciation in British English (1A Synonyms – item 3). Riverside Insights have confirmed that it is acceptable for the assessor to present such items with British pronunciation, as the scoring is based on knowledge, not pronunciation. Similarly, for one item in General Information (8A Where Item 5) the word generally used in Britain is different from the one used in the US. Riverside Insights have confirmed that it is acceptable for the assessor to substitute the British word. • Some items in General Information can have a different meaning in British English (8A Where – item 14, and item 17). Riverside Insights have confirmed that a score can be given if the examinee responds with an answer that relates to the British meaning of the word and shows the examinee correctly understands where to find it.

B) Non-Verbal Abilities

Test 2 Number Series - Examinee is presented with a sequence of numbers with one number missing – has to determine the missing number. 42 items. Measures quantitative reasoning and inductive reasoning.

Test 9 Concept Formation - Examinee has to work out a set of basic rules (relating to visual features) why some items are in a box whereas others are not. Examples are given to demonstrate what is required, and examinee receives immediate feedback regarding correctness of each response before a new item is presented, thus providing a controlled learning task. 40 items. Measures categorical reasoning based on principles of inductive logic.

Number Series and Concept Formation form the Fluid reasoning cluster (Gf). Gf = broad ability to reason, form concepts, and solve problems using unfamiliar information or novel procedures. Gf includes:

- Inductive and deductive reasoning
- Identifying relations,
- Drawing inferences,
- Recognising and forming concepts,
- Identifying conjunctions,
- Recognising disjunctions.

Needs deliberate and flexible control of attention to solve on the spot problems.

Test 7: Visualisation. 2 subtests

- Spatial Relations – examinee has to identify 2 or 3 pieces that form a complete target shape (though may be flipped, rotated).
- Block Rotation – examinee has to identify 2 patterns of blocks that match the target pattern.

It is necessary to administer both tests to achieve a Visualisation score. Visualisation forms part of the Visual Processing cluster. It measures ability to perceive, analyse, synthesize and think with visual patterns. It includes:

- Ability to store and recall visual representations
- Ability to manipulate objects or patterns mentally
- Ability to identify visual representations that appear in vague or obscure circumstances
- Visual imagery
- Visual memory

Strengths	Issues to consider
• WJIV Cog breaks non-verbal reasoning down into a series of tasks each of which measures a specific aspect of visual reasoning. For the purposes of a typical diagnostic assessment Concept Formation and Number Series together would provide sufficient evidence of the non-verbal reasoning abilities of an examinee. They contribute to the Gf-Gc Composite which (as noted in the Overview of Strengths above) enables useful evaluation of discrepancies within the cognitive profile. • Concept Formation is a controlled learning task which measures the examinee's ability to absorb and apply learning from basic principles. It involves only very basic language knowledge is not dependent upon the type of maths training received. These factors make it a useful tool for evaluating the reasoning skills of any individual whose language knowledge or maths knowledge may be underdeveloped (due to a specific learning difficulty or EAL) • Observation of strategies used and errors made in the Number Series test can give some really useful insights into the accuracy, extent and automaticity of the examinee's knowledge of number facts.	• Lacks a construction task • Concept Formation - assessors need to be very familiar with the logic that underpins this test before administering it. • The Fluid Reasoning Cluster contains one other test: Test 15 Analysis-Synthesis. In the interests of not over-testing, it is suggested Analysis Synthesis is only administered if it contributes additional insights into an examinee's profile that add something meaningful to making a diagnosis. Analysis-Synthesis is quite challenging and time consuming to administer. For the examinee it makes heavy demands on working memory and visual sequential ordering. • Visualisation is not in the Fluid Reasoning cluster. It is the test which most closely resembles matrix and design type tests in other intelligence batteries. WJIV regards Visualisation as a Gv task rather than a Gf task – "*Typical GV tasks include recognising rotations and reversal of figures, finding hidden figures, identifying incomplete or distorted figures, and comprehending spatial configurations.*" The tests are quite time consuming to administer, and may only be necessary where the assessor suspects visual perception and spatial reasoning may be one of the examinee's main areas of strength or resilience or may be an area of particular risk (contributing to difficulties in mathematics or problems with understanding visually presented data).

C) Working Memory

Test 3 Verbal Attention - Examinee listens to an intermingled series of animals and digits. Rather than repeating the series or regrouping the items into animals and digits, the examinee answers specific questions (eg the animal that came before the 5). Measures auditory working memory. Taps attentional control – critical to efficient working memory

Test 10: Numbers Reversed - Examinee holds a span of numbers in memory while reversing the sequence of numbers. Measures short-term auditory working memory, and in particular working memory capacity.

These tests form part of the Short-term Working Memory Cluster (Gwm). Gwm is the ability to apprehend and hold information in immediate awareness and then use or manipulate it to carry out a goal. Info is typically retained only a few seconds before it is lost or transferred.

Strengths	Issues to consider
• The combination of Verbal Attention and Numbers Reversed gives sufficient indication of an examinee's auditory working memory for the purposes of a diagnostic assessment. • In Verbal Attention, the fact that an instruction is given after the sequence has been presented, makes it difficult to use strategies to boost scores. It seems to be a very effective way of measuring the working memory's ability to cope in a typical learning situation. • Numbers Reversed is a standard digit span test of working memory. • Short-term Working Memory is one of the clusters that can be compared against the Gf-Gc Composite	• There is a further test in the Short-term Working Memory Cluster – Object Number Sequencing. In the interests of not over-testing, it is suggested that Object Number Sequencing is only administered if it contributes additional insights into an examinee's profile that add something meaningful to making a diagnosis. For example, the manual suggests that a comparison of performance in Verbal Attention and Object Number Sequencing offers the potential to differentiate between limited memory capacity and inattention. • All the tests of auditory working memory include numbers which makes it difficult to compare performance in numbers versus performance without numbers. However qualitative judgements could be made about relative ability to remember numbers versus objects in Verbal Attention.

D) Phonological Processing

Test 5: Phonological Processing

- Words Access – examinee provides a word that has a specific phoneme (at beginning, in middle, or at end).
- Word Fluency – examinee names as many words as possible that begin with a specified sound (two items, each timed for 1 minute)
- Substitution – examinee substitutes part of a word to create a new word. Original word is given in an audio recording.

It is necessary to administer all three of these subtests to achieve a Phonological Processing score. The tests measure complex auditory processing tasks that include speed of lexical access (a narrow ability of long-term retrieval).

Test 12: Nonword Repetition - Examinee repeats nonsense words presented on audio recording. Cannot replay items. Cognitively complex measure of phonological processing, measuring aspects of auditory processing and short-term working memory.

Phonological Processing and Nonword Repetition form the Auditory Processing cluster (GA). GA is the Ability to encode, synthesize and discriminate auditory stimuli, including ability to employ auditory information in task performance.

Strengths	Issues to consider
• The Auditory Processing cluster contains a range of phonological tasks involving changing sounds within words, producing words containing specific sounds, and remembering sequences of sounds. • Words Access, Word Fluency and Substitution measure aspects of phonological awareness. • Word Access and Word Fluency are much more effective at picking up phonological processing difficulties than might be expected given the apparent simplicity of the tasks. Examinees can find it difficult to generate words even if they have an extensive vocabulary knowledge. The timed element of Word Fluency adds in speed of lexical access – making it a good indicator of word retrieval problems. Auditory Processing is one of the clusters that	• The Word Access and Substitution tests are less extensive than equivalent tests of phonological awareness in CTOPP2 and PHAB2. In particular Substitution is not particularly stretching for many older children and adults. • No test of Rapid Naming • There are some issues regarding differences in American and British pronunciation for some of the sounds in the Word Access and Substitution tests. • Word Access and Word Fluency require the examinee to generate words. In some cases underperformance could be attributable to EAL and/or underdeveloped language knowledge rather than weaknesses in phonological processing. • In Substitution there is at least one item where the American

can be compared against the Gf-Gc Composite. • Nonword Repetition reflects a wide range of phoneme combinations	pronunciation of the vowel may be different from some British pronunciation. This could cause confusion on the part of the examinee if he/she does not recognise the word on the audio recording, or translates it to the pronunciation of the vowel he/she knows. Riverside Insight have indicated that a score can be given if the examinee produces a substitution that reflects the British pronunciation, as the main goal of the test is to see if the examinee can understand the question and answer it appropriately. If they can complete using vowel sounds that are the same as the word as they know it, they demonstrate they understand and can complete the task.

E) Visual Processing

Test 4: Letter Pattern Matching Examinee locates and draws a line through two identical letter patterns in rows of six letter patterns. 3 minute time limit. Measures speed at which can make visual symbol discriminations and identify common orthographic (spelling) patterns.

Test 17: Pair Cancellation - Examinee is asked to locate and circle a repeated pattern as quickly as possible. 3 minute time limit. Provides information about

- interference and inhibition control (executive processing)
- sustained attention (attention/concentration)
- ability to perform a simple cognitive task under time pressure (processing speed)

Letter Pattern Matching and Pair Cancellation form the Cognitive Processing Speed cluster (Gs). Gs is the ability to quickly perform both simple and complex cognitive tasks, particularly when under pressure to sustain controlled attention and concentration.

Strengths	Issues to consider
• Tests in the Cognitive Processing Speed cluster are specifically designed to measure processing speed. • Both tests make relatively few demands on working memory, motor coordination or visual tracking (there is no requirement to look up and down from a key). • Useful comparisons can be made between performance with letters (Letter Pattern Matching) and with pictures (Pair Cancellation). • Cognitive Processing Speed is one of the clusters that can be compared against the Gf-Gc Composite	• Pair Cancellation layout may be problematical for some examinees with visual discomfort/visual difficulties.

REMANING SUBTESTS

Test 6: Story Recall

Examinee recalls as many details as possible from increasingly complex stories presented from an audio recording. Measures meaningful memory plus some aspects of oral language development.

Story Recall is one of several tests of memory (see also Visual Auditory Learning, Picture Recognition and Memory for Word) which are not core elements of a typical diagnostic assessment, but could be administered in individual cases if they would add something meaningful to making a diagnosis.

Test 11: Number Pattern Matching - Examinee locates and draws a line through two identical numbers in a row of 6 numbers – difficulty increases from single-digit numbers to triple-digit numbers. 3 minute time limit. Measures the speed at which an examinee can make visual symbol discriminations.

Number Pattern Matching is very similar in format to Letter Pattern Matching, and it would only be useful to administer both tests if the assessor is concerned that the results of the Letter Pattern Matching test may have been affected by visual or EAL related problems with recognising letters, and wishes to be able to compare the examinee's performance with numbers.

Test 13: Visual Auditory Learning - Examinee learns, stores and retrieves novel pictographic representations of words (sometimes a line drawing that relates to the word, sometimes a completely unrelated symbol) to "read" a story. The number of symbols and length of stories builds up progressively. Examinee receives feedback on responses, which makes this a controlled learning task. Measures long-term storage and retrieval.

Visual Auditory Learning would not normally be a core element of a diagnostic assessment. It could be useful where an examinee is having early word reading difficulties, to establish how easily that individual can learn, retain and apply new symbols and their meaning. It would also possibly be a useful indicator of learning potential in a new context (eg with EAL examinees).

Test 14 Picture Recognition
Examinee is required to identify previously presented pictures (seen for 5 seconds) from a larger collection of pictures of the same type of object (eg several different windows). Same type of object is used to eliminate using words to help with recall. Measures visual memory for objects or pictures.

Tests of visual memory are not a core element of a diagnostic assessment, but this test could be administered if identification of visual memory problems would add something meaning to a diagnosis.

Test 15: Analysis-Synthesis - A controlled learning task designed to measure the ability to reason and draw conclusions from given conditions. Uses a directional key to decide what colour a blank square should be. Includes a colour pre-test to check that the examinee can discriminate and name colours accurately. Primarily measures sequential (deductive) reasoning.

In the interests of not over-testing it is suggested that Analysis-Synthesis is only administered if it contributes additional insights into an examinee's profile that add something meaningful to making a diagnosis. It is quite challenging and time consuming to administer. For the examinee it makes heavy demands in working memory and visual sequential ordering – solving the problems requires retention of several reasoning steps in the working memory.

Test 16: Object Number Sequencing - Examinee listens to series of digits and words, then reorders the information with objects first in sequential order and then digits in sequential order. Measures short-term auditory working memory, and in particular working memory capacity. Ability to hold information in immediate awareness, divide the information into 2 groups, and shift attentional resources to the two new ordered sequences.

In the interests of not over-testing it is suggested that Object Number Sequencing is only administered if it contributes additional insights into an examinee's profile that add something meaningful to making a diagnosis. For example, the manual suggests that a comparison of performance in Verbal Attention an Object Number Sequencing offers the potential to differentiate between limited memory capacity and inattention.

Test 18 Memory for Words Examinee repeats lists of unrelated words in the correct sequence. Measures short-term auditory working memory span.

As noted in the Working Memory section above, the combination of Verbal Attention and Numbers Reversed gives sufficient indication of an examinee's auditory working memory for the purposes of a diagnostic assessment.