

An Independent Analysis of the RIOT IQ Test

Christopher G. Jacobs

November 8, 2026

Independent researcher; unaffiliated with RIOT IQ
DOI: 10.5281/zenodo.21972420

Contact: cgjacobs@protonmail.com

Contents

1 Introduction **1**

2 Test Overview **2**

3 Methodology **4**
3.1 Raw subtest scores . . . 4
3.2 Standardization . . . 4
3.3 Composite scores . . . 4
3.4 Composite reliability and error margin . . . 5
3.5 FSIQ eligibility . . . 5

4 Norms **6**

5 Reliability **7**

6 Intercorrelations **8**
6.1 Observed Intercorrelations . . . 8
6.2 Heatmap and Clustering Patterns . . . 8
6.3 Multidimensional Scaling . . . 10

7 Exploratory Factor Analysis **11**
7.1 Number of Factors . . . 11
7.2 Four-Factor Solution . . . 12
7.3 Six-Factor Solution . . . 13
7.4 EFA Discussion . . . 15

8 Confirmatory Factor Analysis **18**
8.1 Model Reproduction . . . 18
8.2 Competing Models . . . 21
8.3 CFA Discussion . . . 25

9 Subtest and Composite Loadings **27**

10 Recommendations **28**

11 Conclusion **32**

A Raw-to-T-Score Lookup Tables **34**
A.1 Raw-to-T-score conversion, VRI, FRI, and SAI subtests, Age 18–39 . . . 34
A.2 Raw-to-T-score conversion, VRI, FRI, and SAI subtests, Age 40–59 . . . 35
A.3 Raw-to-T-score conversion, VRI, FRI, and SAI subtests, Age 60+ . . . 36
A.4 Raw-to-T-score conversion, WMI subtests, by age group . . . 37
A.5 Raw-to-T-score conversion, PSI subtests, by age group . . . 38
A.6 Raw-to-T-score conversion, RTI subtests, by age group . . . 39

B T-Score to Composite Lookup Tables **40**
B.1 ST to FSIQ lookup table – Basic Test . . . 41
B.2 ST to FSIQ lookup table – Full Test . . . 47
B.3 ST to index composite lookup table . . . 51

C Full `lavaan` Model Output **53**

1 Introduction

The Reasoning and Intelligence Online Test (RIOT), developed by psychologist Russell T. Warne and his team, is a commercially available online intelligence test that launched on July 18, 2025 (Warne, 2025f). It reports a Full Scale IQ (FSIQ) composite and six CHC-grounded index scores – Verbal Reasoning, Fluid Reasoning, Spatial Ability, Working Memory, Processing Speed, and Reaction Time – across 15 subtests (Warne, 2025b, 2025d). RIOT markets itself not as "just another" online IQ test but as one that meets the standards of traditional, individually administered batteries, and it bills itself as the "world's best online IQ test" (Warne, 2025b).

These claims rest on an unusually strong public commitment to transparency: RIOT pledged to freely publish its data and analysis code so outside experts could audit the test, aiming to be "the most transparent test ever created" (Warne, 2025c) and "completely trustworthy and auditable" (Warne, 2025e). Its own articles frequently claim that a test should not be considered legitimate without documented psychometric properties and independent evaluation (Warne, 2025g).

More than a year after launch, none of that has materialized. The only public technical documentation is a six-page Informational Bulletin reporting summary-level reliability, factor-analytic, and criterion-validity statistics, without any independent analysis, data, or code (Warne, 2025a). The full manual – reportedly complete at 100+ pages – remains gated behind a request-access process rather than openly released (Warne, n.d.), in direct conflict with the original promise of public access (Warne, 2025c, 2025e). The lone peer-reviewed study using RIOT data addresses response-speed modeling rather than validating RIOT scores themselves (Warne, 2026). A test marketed as the "best online IQ test" in the world, and pledging that IQ tests should not be a "black box," has nonetheless gone entirely unverified by anyone other than its own creators.

This report undertakes that independent verification by looking not just at RIOT's sole published Informational Bulletin (Warne, 2025a), but also at the client-side code that RIOT serves to the browser, from which the norm tables, observed subtest intercorrelations, and factor loadings can be reconstructed. It has three aims: (1) reproduce the developer's reported statistics, most notably the hierarchical six-factor confirmatory model; (2) fit independent exploratory and confirmatory analyses, testing alternative factor structures, to evaluate how well the six-index architecture holds up on its own merits; and (3) assess the reported norming, reliability, and intercorrelation evidence against the standards RIOT sets for itself. Because individual-level data are unavailable, however, the report cannot independently verify the norms, examine test-retest reliability, or cross-validate RIOT against other established measures.

2 Test Overview

The Reasoning and Intelligence Online Test (RIOT) is composed of 15 subtests organized into six broad ability indices: Verbal Reasoning, Fluid Reasoning, Spatial Ability, Working Memory, Processing Speed, and Reaction Time (Warne, 2025b). According to Warne (2025d), these subtests were selected from an initial pool of 37 candidates considered by the test developer, evaluated on criteria including interface requirements, portability across devices, estimated *g*-loading based on prior research, ease of item construction, and response format. Most indices are measured by three subtests each; Processing Speed is measured by two, and Reaction Time is measured by a single subtest with two timed portions (simple and choice reaction time), yielding 16 subscores in total (Warne, 2025d).

Table 1 describes the six indices conceptually, focusing on the ability each is intended to measure and its corresponding broad ability in the Cattell-Horn-Carroll (CHC) taxonomy (Warne, 2025a, 2025d). Table 2 then lists each subtest, its abbreviation (consistent with Table 6), and the index it is placed under.

Table 1: RIOT indices and their conceptual basis

Index	Abbr.	CHC ability	Conceptual description
Verbal Reasoning	VRI	Gc (Comprehension-Knowledge)	Acquired verbal knowledge and the ability to reason with previously learned, culturally transmitted information.
Fluid Reasoning	FRI	Gf (Fluid Reasoning)	Novel problem solving and inductive/deductive reasoning largely independent of acquired knowledge.
Spatial Ability	SAI	Gv (Visual-Spatial Processing)	Generating, storing, retrieving, and mentally manipulating visual images and spatial relationships.
Working Memory	WMI	Gwm (Working Memory)	Holding and mentally manipulating a limited amount of information over a brief period.
Processing Speed	PSI	Gs (Processing Speed)	Performing simple, well-learned cognitive tasks quickly and accurately under time pressure.
Reaction Time	RTI	Gt (Reaction and Decision Speed)	Speed of reacting to and making simple decisions about a stimulus, reflecting basic neural processing speed.

Some factor models below use VSI (Visual-Spatial Index) in place of SAI. VSI is the more common label for this ability in the psychometric literature, and SAI was inadvertently labeled VSI in parts of the confirmatory factor analysis. The two labels are interchangeable and refer to the same Gv factor.

Table 2: RIOT subtests by index

Subtest	Abbr.	Index	Description
Vocabulary	VO	VRI	Multiple-choice identification of the word closest in meaning to a target word.
Information	IN	VRI	Multiple-choice general-knowledge questions spanning domains such as history, science, and geography.
Analogies	AN	VRI	Multiple-choice word-pair analogies (A is to B as C is to ?).
Matrix Reasoning	MR	FRI	Identifying the image that completes the pattern in a 3×3 grid.
Visual Puzzles	VP	FRI	Selecting which 2–4 of eight pieces assemble into a target image.
Figure Weights	FW	FRI	Using balance-scale stimuli to infer which option balances a final scale.
Object Rotation	OR	SAI	Identifying a target shape after it has been rotated in two or three dimensions.
SToVeS	STV	SAI	Verbally presented items requiring reasoning about cardinal directions and turns.
Spatial Orientation	SO	SAI	Map-based items requiring navigation, perspective-taking, or locating landmarks.
Computation Span	CS	WMI	Recalling, in order, the answers to a sequence of simple arithmetic problems.
Exposure Memory	EM	WMI	Recognizing which images were shown in a briefly presented sequence.
Visual Reversal	VR	WMI	Recalling, in reverse order, a sequence of color changes on a 3×3 grid.
Symbol Search	SS	PSI	Speeded (2-minute) task judging whether either of two target symbols appears in a set of five.
Abstract Matching	AM	PSI	Speeded task selecting which of two options more closely matches a target shape varying on shape, color, number, and orientation.
Simple Reaction Time	SRT	RTI	Pressing a key as soon as a single stimulus appears; 20 trials.
Choice Reaction Time	CRT	RTI	Pressing one of two keys depending on which of two stimuli appears; 20 trials.

Note. Simple Reaction Time and Choice Reaction Time are two timed portions of a single Reaction Time subtest, each yielding a separate subscore.

Administration format varies by subtest. Most subtests are self-paced and can be completed in any order, with breaks permitted between subtests. According to the developer's technical bulletin, the full battery is typically administered in 75 to 100 minutes, including instructions (Warne, 2025a).

3 Methodology

This section outlines the scoring methodology used by RIOT, including how raw subtest scores are calculated, how they are standardized into z-scores and T-scores, how composite scores are computed from subtest scores, and how reliability and error margins are derived (Warne, 2025b).

3.1 Raw subtest scores

For most subtests, the raw score is simply the sum of item scores. However, three subtests are scored differently. In the two formulas below, t_i is the response time recorded for item i and n_c is the number of items answered correctly.

- **Reaction Time.** The simple and choice portions are scored separately. Each raw score is the mean response time across valid trials – the first 5 of the 20 administered trials are treated as practice and discarded, and a trial counts as valid only if the response was correct and fell between 150 ms and 10,000 ms. The score is negated because a lower reaction time indicates better performance.
- **Symbol Search.** Items are filtered to correct responses only; the raw score is the negative ratio of response time summed over those correct items to the number correct:

$$x_{\text{SS}} = -\frac{\sum_{i \in \text{correct}} t_i}{n_c}.$$

- **Abstract Matching.** The raw score is a similar construction – a negative ratio of response time to number correct – but the numerator is response time summed over every item attempted, rather than over correct items only as in Symbol Search:

$$x_{\text{AM}} = -\frac{\sum_{i \in \text{all items}} t_i}{n_c}.$$

  Time spent on wrong answers therefore drags the score down here in a way it cannot in Symbol Search.

3.2 Standardization

Each raw score x_i on subtest i is converted to a z-score against the raw-score norms in Table 3, using the mean and SD for the examinee's age bracket a at test completion – 18–39, 40–59, or 60+:

$$z_i = \frac{x_i - \mu_{i,a}}{\sigma_{i,a}}.$$

A subtest or index reported alone is scaled directly from its own z_i using the $T = 50 + 10z$ transformation introduced in the Norms section.

3.3 Composite scores

Index scores (e.g., Fluid Reasoning) and the FSIQ are both composites – weighted combinations of the z-scores of the subtests that feed them, restricted to whichever of those subtests the examinee actually took. Writing C for that subtest set, the composite z-score is

$$Z_C = \frac{\sum_{i \in C} \lambda_i z_i}{\sigma_C}, \qquad \sigma_C = \sqrt{\sum_{i \in C} \lambda_i^2 + \sum_{i \in C} \sum_{\substack{j \in C \\ j \neq i}} \lambda_i \lambda_j r_{ij}},$$

where λ_i is subtest i's weight and r_{ij} its correlation with subtest j (Table 6). Notably, the weight λ_i RIOT uses here is exactly the "Factor loading" value reported per subtest in Table 5 – i.e., composites are explicit loading-weighted sums of standardized subtest scores, not simple averages. The denominator σ_C is the standard deviation of that weighted sum under the observed subtest intercorrelations; dividing by it renormalizes Z_C back onto a standard-normal scale regardless of how many, or which, subtests happen to be in C. This is what lets the same formula score a full six-index battery, a partial Custom test, or an index missing one of its usual subtests, without RIOT needing to store separate weights for every possible subtest combination.

Composite Z_C is then rescaled the same way as individual subtests, clipped to a hard floor/ceiling:

$$\text{Score}_C = \min\Big(\max\big(\text{center} + \text{factor} \cdot Z_C,\ \min\big),\ \max\Big),$$

with $(\text{center}, \text{factor}, \min, \max) = (100, 15, 71, 145)$ for the FSIQ and $(50, 10, 31, 80)$ for the six indices (and for individual subtests).

These artificial floors and ceilings are highly questionable and conflict directly with the composite effect: because the subtests are only imperfectly correlated, scoring high across all of them is rarer than scoring high on any one, so an extreme composite is rarer than extreme scores on its parts and needs a *wider* range, not an identical one. The WAIS does exactly this: its subtests are capped at a scaled score of 19 (+3.0 SD), but its FSIQ ceiling is 160 (+4.0 SD) (Wechsler, 2008; Wechsler et al., 2024). RIOT instead caps both the subtest and composite ceilings at +3.0 SD, discarding valuable information and making all examinees above 145 (or below 71) indistinguishable from one another; Appendix B recovers the true ceilings and floors for the FSIQ and index composites.

3.4 Composite reliability and error margin

Composite reliability follows Mosier's classical formula for the reliability of a weighted composite of correlated components (Mosier, 1943):

$$r_{cc'} = 1 - \frac{\sum_{i \in C} \lambda_i^2 \big(1 - r_{xx,i}\big)}{\sigma_C^2},$$

reusing the same σ_C^2 from the composite-score formula above and each subtest's individual reliability $r_{xx,i}$ (Table 5). For the six indices and for the Basic and Full product-level FSIQs, the test caches this result rather than recomputing it at score time; those cached values agree to two decimal places with values recomputed from Mosier's formula using Tables 5 and 6. Custom tests, which can combine an arbitrary subtest set, fall back to computing $r_{cc'}$ directly from the formula.

The reported margin of error is a 95% confidence interval half-width built from this reliability, using the same (center, factor) scaling as the composite score itself:

$$\text{ME}_{95} = 1.96 \times \text{factor} \times \sqrt{1 - r_{cc'}}.$$

This is the ± figure (e.g., "±3.5 IQ") attached to each product and quoted in RIOT's marketing.

3.5 FSIQ eligibility

An FSIQ is not always reported. It is reported only if at least one Verbal Reasoning, one Fluid Reasoning, and one Spatial Ability subtest were taken, and those three indices together account for at least half of all distinct subtests administered (Simple and Choice Reaction Time count as a single "Reaction Time" subtest for this purpose).

4 Norms

RIOT's norm sample consists of 417 individuals, which the RIOT team describes as nationally representative of the test's target population (English-speaking adults born in the United States) in terms of sex, age, race/ethnicity, and education level. Some analyses are supplemented with data from a separate 1,620-person Stage 3 tryout sample from the same population (Warne, 2025a). However, beyond this summary, the bulletin does not disclose any demographic breakdown, recruitment method, or stratification procedures through which that representativeness was verified. Table 3 reports the raw-score norms (mean and SD) for each standard-length subtest across the three age brackets used in norming (18–39, 40–59, and 60+).

To make the raw-score norms in Table 3 more directly interpretable, Appendix A converts them into raw-to-T-score lookup tables for all sixteen subtests, using the standard $T = 50 + 10\left(\frac{\text{raw}-\text{mean}}{\text{SD}}\right)$ transformation. The nine Verbal Reasoning, Fluid Reasoning, and Spatial Ability subtests are each tabulated in one table per age bracket (Tables A.1, A.2, and A.3), while the Working Memory (Table A.4), Processing Speed (Table A.5), and Reaction Time (Table A.6) subtests are each tabulated in a single table spanning all three age brackets.

Because the number of items per subtest is unknown, the true raw-score ceiling of each subtest is also unknown. However, because RIOT imposes a hard 31T floor and 80T ceiling, each table runs only up to the raw score at which the formula first reaches 80T.

Table 3: Standard subtest raw-score norms by age group

		18–39		40–59		60+	
Subtest	Index	Mean	SD	Mean	SD	Mean	SD
Vocabulary	VRI	12.18	4.41	13.33	5.12	14.74	5.77
Information	VRI	9.65	5.15	10.77	5.51	12.02	5.96
Analogies	VRI	12.35	6.16	13.21	6.68	14.61	7.16
Matrix Reasoning	FRI	15.14	5.11	13.80	5.23	12.79	5.18
Visual Puzzles	FRI	14.82	5.81	15.15	5.65	14.47	4.83
Figure Weights	FRI	11.10	4.06	10.75	3.88	9.90	3.65
Object Rotation	SAI	14.17	4.36	13.79	3.76	12.73	3.66
SToVeS	SAI	8.09	4.87	7.13	4.43	7.33	4.48
Spatial Orientation	SAI	10.69	5.40	10.36	5.11	10.23	5.21
Computation Span	WMI	7.31	2.78	7.47	3.02	7.14	2.66
Exposure Memory	WMI	6.15	3.56	5.57	3.09	4.17	2.25
Visual Reversal	WMI	7.22	2.85	7.16	2.74	6.85	2.23
Symbol Search	PSI	3.41s	1.19s	3.47s	0.95s	3.84s	0.99s
Abstract Matching	PSI	1.90s	0.43s	2.07s	0.53s	2.24s	0.58s
Simple Reaction Time	RTI	370ms	100ms	400ms	120ms	430ms	120ms
Choice Reaction Time	RTI	530ms	140ms	570ms	140ms	590ms	160ms

5 Reliability

RIOT's published reliability evidence is limited to internal consistency. Table 4 reports coefficient alpha and omega for the overall IQ score and each of the six indices, as given in the bulletin. No test-retest (temporal stability) reliability has been published for RIOT.

Table 4: Internal consistency reliability (coefficient alpha and omega)

Score	α	ω
Overall IQ	.980	.984
Verbal Reasoning (VRI)	.962	.969
Fluid Reasoning (FRI)	.945	.951
Spatial Ability (SAI)	.943	.951
Working Memory (WMI)	.858	.907
Processing Speed (PSI)	.935	.974
Reaction Time (RTI)	.938	.952

Note. Source: Warne (2025a).

Table 5 reports each subtest's factor loading on its index (from the developer's six-factor model) alongside its individual reliability coefficient (r_{xx}). These are the same subtest-level coefficients used internally, via Mosier's formula, to compute the composite-score reliabilities and margins of error described in the Methodology section. The individual subtests' α and ω coefficients are not published; the bulletin reports only their ranges (α = .766–.936; ω = .847–.968).

Table 5: Subtest factor loadings and reliability

Subtest	Abbr.	Index	Factor loading (λ)	Subtest reliability (r_{xx})
Vocabulary	VO	VRI	0.63	0.92
Information	IN	VRI	0.64	0.94
Analogies	AN	VRI	0.61	0.95
Matrix Reasoning	MR	FRI	0.69	0.91
Visual Puzzles	VP	FRI	0.66	0.90
Figure Weights	FW	FRI	0.72	0.88
Object Rotation	OR	SAI	0.66	0.85
SToVeS	STV	SAI	0.68	0.92
Spatial Orientation	SO	SAI	0.83	0.92
Computation Span	CS	WMI	0.46	0.84
Exposure Memory	EM	WMI	0.41	0.87
Visual Reversal	VR	WMI	0.53	0.86
Symbol Search	SS	PSI	0.52	0.97
Abstract Matching	AM	PSI	0.36	0.95
Simple Reaction Time	SRT	RTI	0.50	0.90
Choice Reaction Time	CRT	RTI	0.49	0.91

The factor loadings are used to weight subtests in the composite score calculations, while the reliability coefficients are used to compute the associated confidence intervals.

6 Intercorrelations

6.1 Observed Intercorrelations

Table 6 presents RIOT's observed inter-subtest correlation matrix. This matrix underlies the composite FSIQ and index scores, composite reliability, and error-margin calculations, all described in the Methodology section. Every one of the 120 pairwise correlations is positive, consistent with the positive manifold expected of a measure of general intelligence.

This intercorrelation matrix also serves as the basis for the exploratory and confirmatory factor analyses later in this report, which test whether the hierarchical structure RIOT proposes in its bulletin is supported by the data, or whether alternative structures fit the observed correlations better.

The matrix recovered from RIOT's client-side code matches the range the bulletin reports for its "development sample" ($r = .099$ to $.613$; Warne, 2025a) to three decimal places, excluding the within-subtest Simple–Choice Reaction Time correlation ($r = .73$). However, it is unknown whether the correlations were computed on raw scores pooled across age groups or on age-standardized scaled scores, as the bulletin does not specify. All analyses in this report assume the intercorrelation matrix is based on the $N = 417$ norm sample, as supported by the close match between our reproduced χ^2 and the bulletin's (Section 8.1). Because χ^2 scales with sample size, the norm and Stage 3 tryout samples combined ($N \approx 2{,}000$) would give a χ^2 roughly five times larger.

Table 6: RIOT intercorrelation matrix

	VO	IN	AN	MR	VP	FW	OR	STV	SO	CS	EM	VR	SS	AM	SRT	CRT
VO	1.00															
IN	.61	1.00														
AN	.59	.55	1.00													
MR	.36	.34	.37	1.00												
VP	.34	.41	.36	.55	1.00											
FW	.44	.40	.38	.61	.54	1.00										
OR	.33	.36	.32	.53	.52	.52	1.00									
STV	.41	.46	.40	.51	.46	.53	.51	1.00								
SO	.55	.58	.52	.55	.56	.55	.58	.61	1.00							
CS	.29	.21	.22	.34	.29	.36	.28	.38	.34	1.00						
EM	.20	.23	.22	.39	.29	.33	.29	.29	.31	.20	1.00					
VR	.25	.18	.24	.40	.42	.43	.41	.36	.44	.32	.24	1.00				
SS	.31	.31	.31	.30	.24	.33	.30	.25	.43	.30	.20	.38	1.00			
AM	.28	.24	.27	.10	.15	.20	.19	.18	.34	.19	.14	.16	.48	1.00		
SRT	.30	.34	.27	.33	.27	.32	.26	.24	.37	.19	.14	.22	.32	.24	1.00	
CRT	.27	.33	.30	.32	.26	.30	.24	.23	.33	.21	.20	.25	.39	.23	.73	1.00

Note. AM = Abstract Matching; AN = Analogies; CRT = Choice Reaction Time; CS = Computation Span; EM = Exposure Memory; FW = Figure Weights; IN = Information; MR = Matrix Reasoning; OR = Object Rotation; SRT = Simple Reaction Time; SO = Spatial Orientation; STV = SToVeS; SS = Symbol Search; VP = Visual Puzzles; VR = Visual Reversal; VO = Vocabulary.

6.2 Heatmap and Clustering Patterns

Even before any formal factor analysis, the clustering visible in Table 6 and Figure 1 gives a sense of what structure to expect. Subtests measuring the same underlying ability should intercorrelate more strongly with one another than with subtests measuring something else.

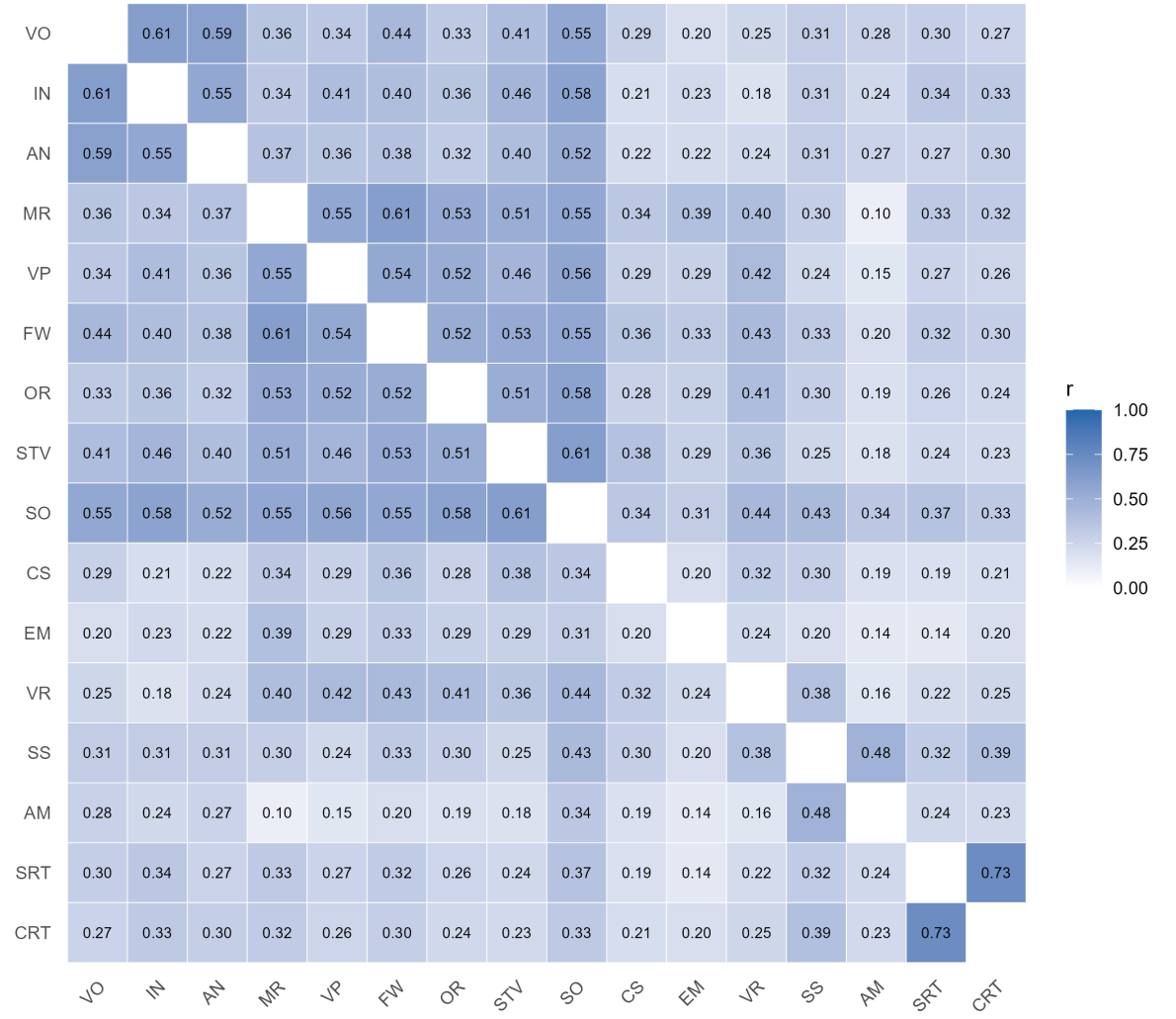


Figure 1: Heatmap of the RIOT subtest intercorrelation matrix

Vocabulary, Information, and Analogies form the clearest such cluster (r = .55–.61), consistent with a single Verbal/Gc factor built on shared crystallized-knowledge demands. Matrix Reasoning, Visual Puzzles, Figure Weights, Object Rotation, SToVeS, and Spatial Orientation form a second, broader cluster: all 15 pairwise correlations among these six subtests fall in a narrow .46–.61 band with no sub-clustering evident, suggesting these subtests may be better described by a single combined factor than by separate Fluid Reasoning and Spatial Ability factors.

Spatial Orientation stands out within that cluster: it has the highest average correlation with the rest of the battery of any subtest ($\bar{r}$ = .47 across the other 15 subtests, compared to .40–.42 for the next-highest, Matrix Reasoning and Figure Weights), and it correlates with Vocabulary, Information, and Analogies (r = .52–.58) nearly as strongly as those three correlate with each other (r = .55–.61) – suggesting Spatial Orientation may be an unusually strong indicator of g itself that also carries a secondary Verbal loading. This is unsurprising for a subtest whose task seems to tap into more than one broad ability at a time.

The three proposed Working Memory subtests, by contrast, do not cluster with one another at all (r = .20–.32 among Computation Span, Exposure Memory, and Visual Reversal) – markedly weaker than either cluster above – and Visual Reversal in particular correlates more strongly with the Fluid/Spatial cluster (r = .36–.44) than with the other Working Memory subtests, foreshadowing possible issues with modeling a Working Memory factor.

Finally, Abstract Matching looks like a very poor indicator of general ability: its correlations with every other subtest are weak (r = .10–.28) except with Symbol Search (r = .48), where it is only moderate.

6.3 Multidimensional Scaling

Figure 2 plots the subtests on a 2D map using multidimensional scaling (MDS), so that subtests that correlate more strongly sit closer together. The map was made with the `smacof` package (de Leeuw & Mair, 2009) in R. The fit is only fair (Stress-1 = .161; Kruskal, 1964), so the overall layout deserves more weight than individual distances. The configuration resembles a radex (Marshalek et al., 1983), with the more g-loaded subtests sitting closer to the center.

The observations from the previous section are all visible here: the Verbal subtests form a tight cluster, the Fluid and Spatial subtests show no separation, the Working Memory subtests are scattered and do not cluster in any meaningful way, and Abstract Matching sits farthest from the center. The subtests closest to the center – Spatial Orientation, followed by Figure Weights and Matrix Reasoning – are also the most g-loaded.

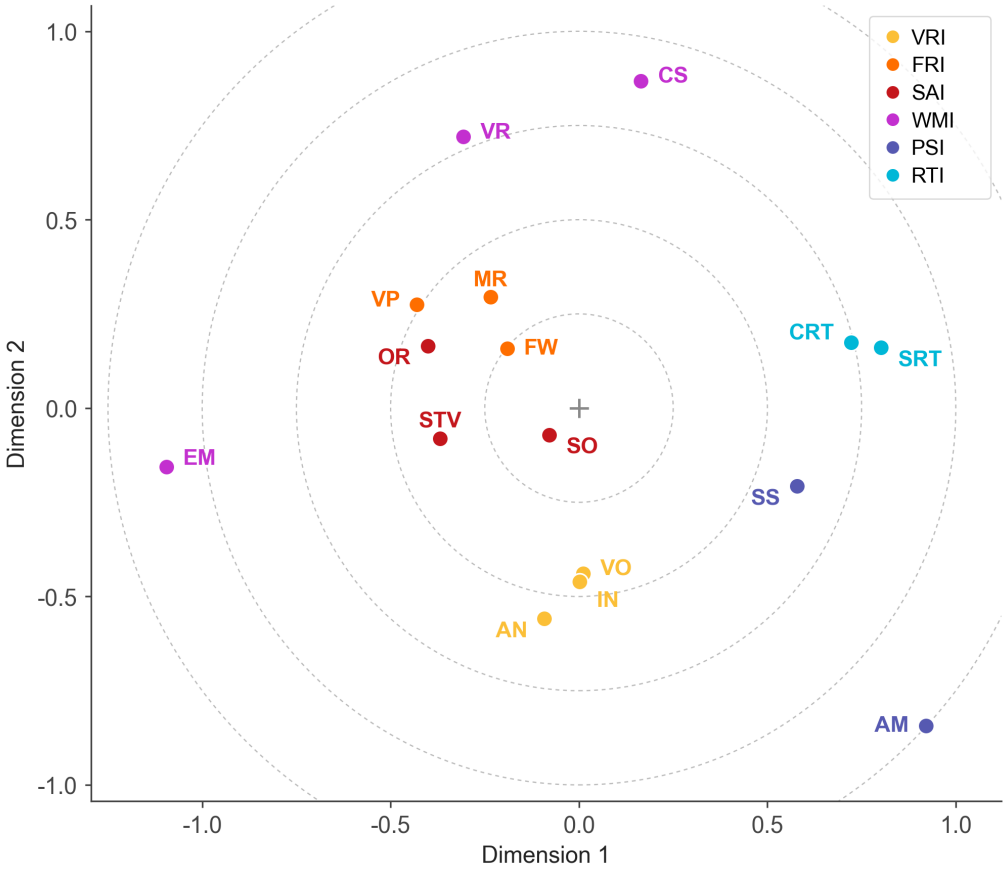


Figure 2: Two-dimensional non-metric MDS of the RIOT subtest intercorrelations

Note. Ordinal SMACOF solution, Stress-1 = .161. Abbreviations as in Table 6.

7 Exploratory Factor Analysis

RIOT's Informational Bulletin (Warne, 2025a) describes the test's underlying structure as a hierarchical six-factor model: each of the 15 subtests (16 subscores, per Table 2) loads onto one of six first-order index factors – Verbal Reasoning, Fluid Reasoning, Spatial Ability, Working Memory, Processing Speed, and Reaction Time – and all six index factors load in turn onto a single second-order general factor (g), which is the basis for the Full Scale IQ (FSIQ) score RIOT ultimately reports.

Before testing this model in a confirmatory framework (Section 8.1), we first examined the subtest intercorrelations (Table 6) using exploratory factor analysis (EFA), which imposes no a priori structure on the data. All analyses were conducted in `R` using the `psych` package (Revelle, 2025), with minimum-residual (minres) extraction and oblimin (oblique) rotation, on the subtest intercorrelation matrix with $N = 417$.

7.1 Number of Factors

The unrotated eigenvalues of the correlation matrix were 6.34, 1.50, 1.23, 1.08, 0.83, 0.77, 0.65, 0.53, . . . , with only the first four exceeding Kaiser's conventional threshold of 1. Parallel analysis (Horn, 1965) agreed, recommending four factors. The very large first eigenvalue (6.34, more than four times the second) indicates a strong general factor underlying the data. Figure 3 depicts the parallel analysis scree plot.

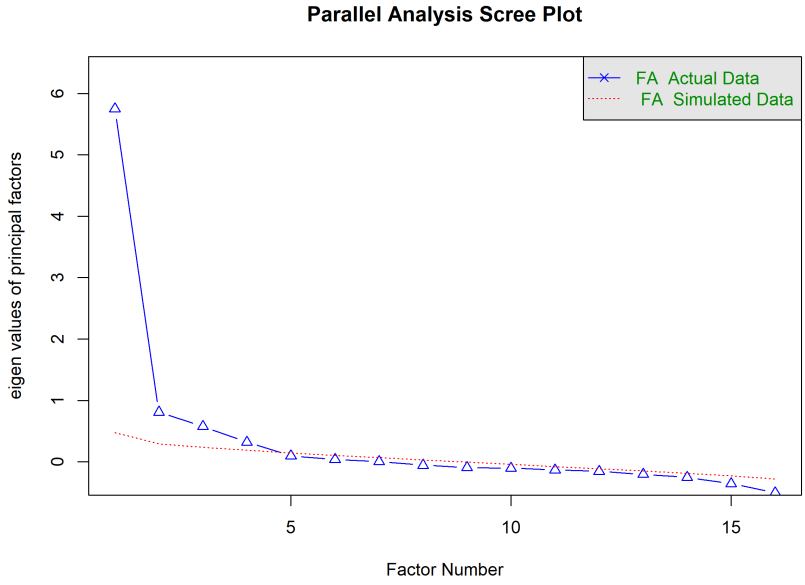


Figure 3: Scree plot and parallel analysis of the RIOT subtest intercorrelation matrix

Because parallel analysis is a purely statistical criterion and our interest here is also in whether the data recover the developer's six-factor model, we fit and examined four-, five-, and six-factor models. Table 7 compares their fit. TLI, RMSEA, and the χ^2 test all improve as more factors are extracted – unsurprising, since added factors mechanically buy fit regardless of whether they capture real structure. The parsimony-penalized BIC, which corrects for this, instead favors the four-factor solution, consistent with the parallel-analysis recommendation. The five-factor solution bears this out directly: its additional factor exists only in name, with no subtest reaching a loading above $|.20|$, only 2% of variance explained, and an ultra-Heywood case flagged during estimation. It was set aside as uninterpretable, lending further support to the four-factor model.

Table 7: Fit of the four-, five-, and six-factor EFA solutions

Factors	χ^2	df	p	TLI	RMSEA	RMSEA 90% CI	BIC
4	100.02	62	.002	.972	.038	[.024, .052]	−274.03
5	68.99	50	.039	.983	.030	[.007, .046]	−232.66
6	52.06	39	.079	.985	.028	[.000, .047]	−183.24

Note. All models: minres extraction, oblimin rotation, $N = 417$.

7.2 Four-Factor Solution

The four-factor solution splits cleanly into a Verbal factor (IN .76, VO .73, AN .64), a Reaction Time factor (CRT .96, SRT .72), a Processing Speed factor (SS .77, AM .56), and a single broad fourth factor combining all nine Fluid Reasoning, Spatial Ability, and Working Memory subtests (MR .80, FW .69, OR .68, VP .68, STV .59, VR .58, SO .48, EM .41, CS .38). As predicted, SO cross-loads moderately on the Verbal factor (.26); VR also shows a smaller cross-loading on the Processing Speed factor (.20).

The solution that both parallel analysis and BIC prefer separates Verbal, Reaction Time, and Processing Speed as distinct factors but does not statistically distinguish Fluid Reasoning, Spatial Ability, and Working Memory from one another. Table 8 reports the corresponding Schmid-Leiman bifactor solution (Schmid & Leiman, 1957), and Figure 4 shows the matching omega bifactor diagram, with the merged Fluid/Spatial/Memory cluster visible as a single group factor alongside g.

Table 8: Four-factor Schmid-Leiman bifactor loadings

Subtest	g	F1	F2	F3	F4	h^2	u^2	p^2	com
VO	0.58	0.51				0.60	0.40	0.56	1.99
IN	0.59	0.53				0.64	0.36	0.55	2.02
AN	0.55	0.45				0.51	0.49	0.60	1.95
SRT	0.48		0.60			0.59	0.41	0.40	1.94
CRT	0.52		0.80			0.92	0.08	0.29	1.71
SS	0.53			0.63		0.68	0.32	0.41	1.97
AM	0.37			0.45		0.36	0.64	0.37	2.20
MR	0.57				0.53	0.61	0.39	0.53	2.07
FW	0.60				0.46	0.57	0.43	0.63	1.89
OR	0.54				0.45	0.50	0.50	0.58	1.96
VP	0.54				0.45	0.51	0.49	0.57	1.98
STV	0.56				0.39	0.51	0.49	0.62	1.99
VR	0.45			0.20	0.38	0.39	0.61	0.51	2.56
SO	0.71	0.26			0.32	0.69	0.31	0.72	1.77
EM	0.34				0.27	0.19	0.81	0.61	1.94
CS	0.39				0.25	0.23	0.77	0.64	2.09

Note. Schmid-Leiman transformation of the oblimin four-factor solution; minres extraction, $N = 417$. Blank cells indicate a loading below |.20|. g = general-factor loading. F1–F4 = residualized group-factor loadings. h^2 = communality; u^2 = uniqueness; p^2 = proportion of communality attributable to g; com = item complexity (Hoffman's index). Abbreviations as in Table 6.

The predicted Fluid/Spatial cluster appears as F4 but is broader than expected: Computation Span, Exposure Memory, and Visual Reversal – the Working Memory subtests – load onto it too, confirming the discriminant-validity failure foreshadowed by their weak within-WMI correlations (r = .20–.32). The "working memory" subtests are effectively orphaned: CS and EM have the weakest loadings in the group (.25–.27), and although VR loads more strongly (.38), the three clearly do not form a coherent Working Memory factor.

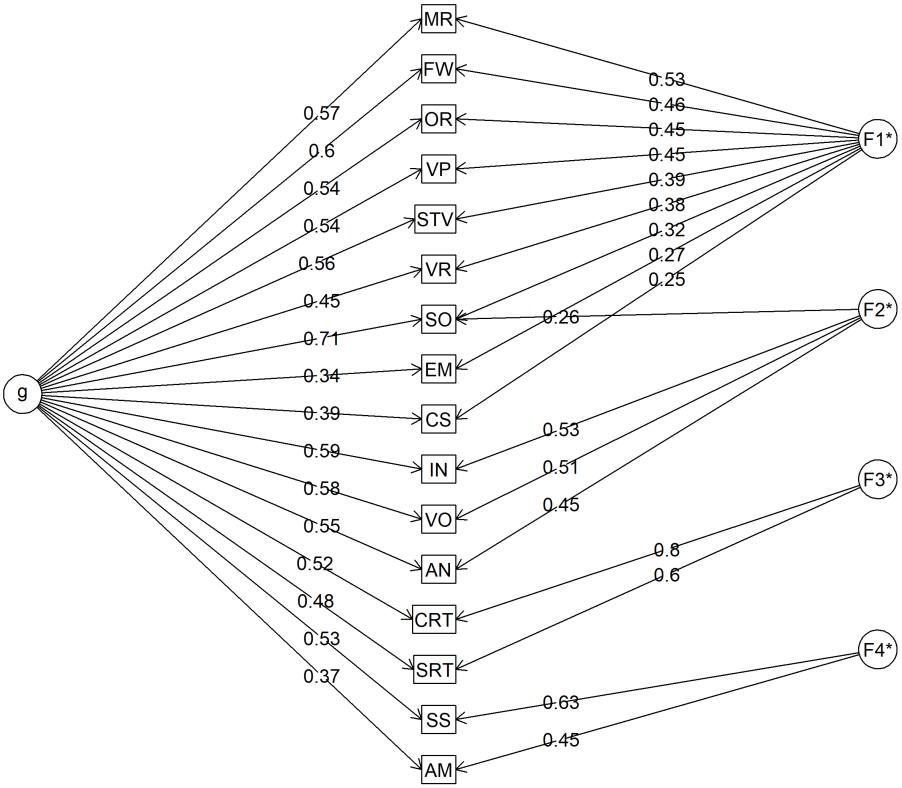


Figure 4: Omega bifactor diagram, four-factor solution

7.3 Six-Factor Solution

We also examine the six-factor extraction in detail because it maps directly onto the developer's a priori six-index architecture and is the only extraction that attempts to separate theorized Fluid Reasoning, Spatial Ability, and Working Memory factors. Table 9 reports the corresponding Schmid-Leiman bifactor solution (Schmid & Leiman, 1957) – each subtest's loading on the general factor (g) alongside its residualized loading on the six group factors, with loadings $\geq |.20|$ shown – which underlies the omega reliability estimates discussed below.

Table 9: Six-factor Schmid-Leiman bifactor loadings

Subtest	g	F1	F2	F3	F4	F5	F6	h^2	u^2	p^2	com
VO	0.58	0.56						0.64	0.36	0.52	2.05
IN	0.57	0.51						0.62	0.38	0.52	2.09
AN	0.54	0.49						0.54	0.47	0.55	2.05
MR	0.65		0.53					0.70	0.30	0.61	1.93
FW	0.66		0.33					0.57	0.43	0.75	1.64
VP	0.59		0.29			0.23		0.51	0.49	0.68	1.83
EM	0.37		0.26					0.21	0.80	0.68	1.90
VR	0.49							0.37	0.63	0.64	2.24
SRT	0.46			0.89				1.00	0.00	0.21	1.50
CRT	0.44			0.57				0.58	0.42	0.34	2.12
SS	0.51				0.76			0.83	0.17	0.31	1.76
AM	0.33				0.39			0.32	0.68	0.33	2.65
SO	0.73	0.22				0.38		0.76	0.24	0.71	1.82
OR	0.59		0.23			0.34		0.54	0.46	0.64	2.00
STV	0.62					0.25		0.54	0.46	0.71	1.84
CS	0.46						0.57	0.54	0.46	0.40	1.94

Note. Schmid-Leiman transformation of the oblimin six-factor solution; minres extraction, $N = 417$. Blank cells indicate a loading below |.20|. g = general-factor loading. F1–F6 = residualized group-factor loadings. h^2 = communality; u^2 = uniqueness; p^2 = proportion of communality attributable to g; com = item complexity (Hoffman's index). Abbreviations as in Table 6.

Every subtest loads substantially on g (range .33 to .73), reaffirming the dominant general factor already evident in the eigenvalues (and in the CFA loadings reported later). After partialling out g, four group factors are recovered cleanly: Verbal (VO, IN, AN), Reaction Time (SRT, CRT), and Processing Speed (SS, AM) retain only their own subtests, and Fluid Reasoning retains its three intended subtests (MR, FW, VP) plus an additional loading from Exposure Memory (.26) – nominally a Working Memory subtest. VP also cross-loads onto Spatial Ability (.23). Spatial Ability recovers all three of its subtests (SO .38, OR .34, STV .25), though Object Rotation cross-loads on Fluid Reasoning (.23) and Spatial Orientation, as expected, cross-loads on Verbal (.22).

Working Memory again is nowhere to be seen: only Computation Span retains a group-factor loading (.57), although it is hard to interpret what this factor actually represents with only a single subtest as an indicator. Exposure Memory sits with Fluid Reasoning (.26), and Visual Reversal has no loading above |.20| on any group factor at all.

Reaction Time is also unstable: Simple Reaction Time is a Heywood case ($h^2 = 1.00$), which is implausible for a subtest with a reliability of .90 (Table 5). SRT also loads well above Choice Reaction Time (.89 vs. .57), whereas the four-factor solution reverses that order (.60 vs. .80; Table 8). Both are typical of a factor with only two indicators, and together with the single-indicator Computation Span factor, this suggests that six factors are more than these data support.

Both solutions therefore indicate that Working Memory is the least supported of the six a priori indices under the Schmid-Leiman decomposition. Figure 5 depicts the full omega bifactor diagram underlying Table 9, showing g alongside all six group factors.

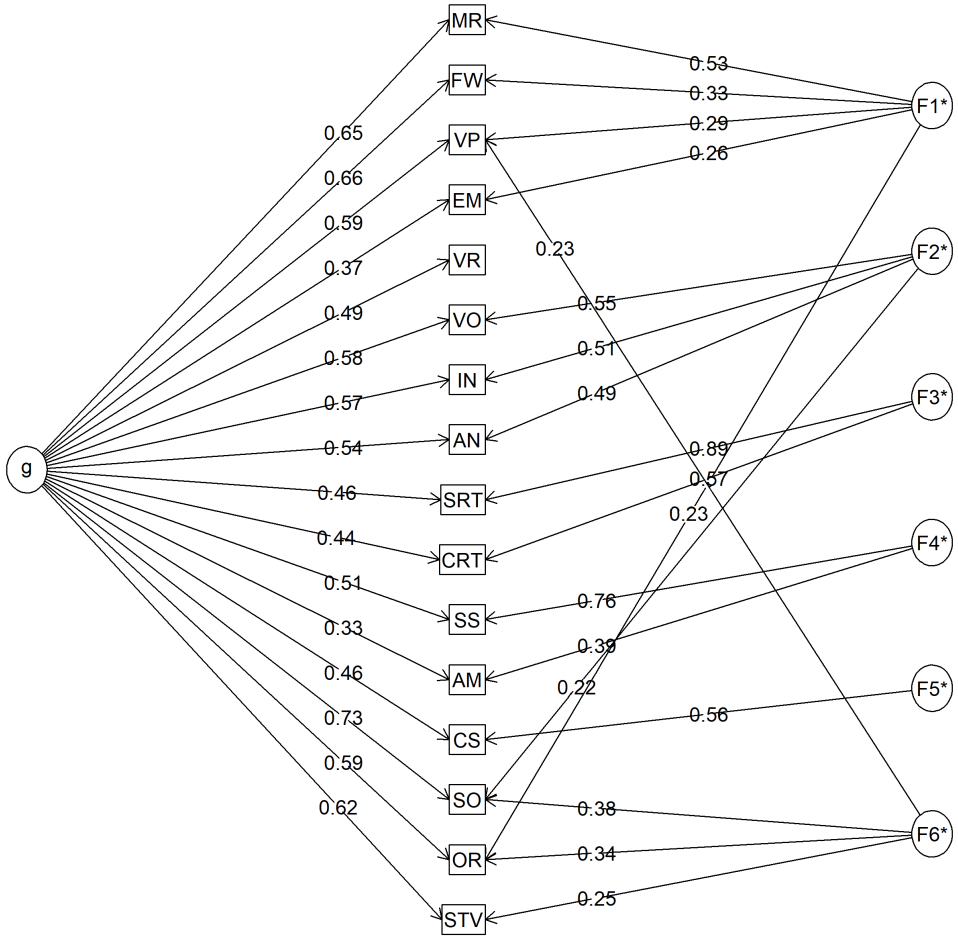


Figure 5: Omega bifactor diagram, six-factor solution

7.4 EFA Discussion

On the evidence assembled in this report, we believe the four-factor solution, not the developer's six-factor architecture, is the better-supported structure for RIOT. It is the solution parallel analysis recommends, and it matches what a plain inspection of the intercorrelation matrix predicts (Section 6.2). Fluid Reasoning and Spatial Ability show no clean internal separation in the raw correlations, while Verbal, Reaction Time, and Processing Speed do.

The attempted five-factor EFA reinforces this. As noted above, when a fifth factor is given room to let Fluid Reasoning and Spatial Ability separate, no coherent split emerges: rather than dividing into FRI and VSI, the additional factor comes back empty, with no subtest reaching a loading above |.20|, only 2% of variance explained, and an ultra-Heywood case flagged during estimation. This suggests the Fluid/Spatial split visible in the six-factor Schmid-Leiman solution (Table 9) is more an artifact of forcing the extraction to six factors than a structure the data organically support. The six-factor solution's other estimation problems – a Heywood case for Simple Reaction Time and a group factor defined by Computation Span alone – point the same way.

This is not to imply that Fluid Reasoning and Spatial Ability are one and the same. CHC theory treats them as distinct broad abilities, and current professional intelligence tests routinely separate them: both the WAIS-V (Wechsler et al., 2024) and WISC-V (Wechsler, 2014) report separate Fluid Reasoning and Visual-Spatial indices. While WAIS-IV did have a combined

Perceptual Reasoning Index (PRI) (Wechsler, 2008), this likely reflected the specific combination of subtests chosen for its 10-subtest core battery.

Weiss et al. (2013) directly tested whether a Gf/Gv split is empirically justified. The WAIS-IV technical manual (Wechsler, 2008) derives its standard four-factor structure from only 10 *core* subtests – the minimum needed to compute the four index scores and FSIQ – while the full battery also includes 5 *supplemental* subtests that are not required for standard scoring but are available for substitution or extended interpretation. Weiss et al. incorporated these five supplemental subtests alongside the 10 core ones, comparing four- and five-factor higher-order models across the full 15-subtest battery, and found the five-factor solution – separating PRI into a three-subtest Visual-Spatial factor (Block Design, Visual Puzzles, Picture Completion) and a three-subtest Fluid Reasoning factor (Matrix Reasoning, Arithmetic, Figure Weights) – fit better than the traditional four-factor PRI structure.

Notably, the original 10-subtest core battery had only one Fluid Reasoning subtest (MR) to work with, too few to support a separately identified factor, which is why Fluid Reasoning merged into Spatial Ability as "Perceptual Reasoning" in the first place; the finding of Weiss et al. (2013) indicates that once enough indicators of each ability are available – three VSI and three FRI, from the full 15-subtest battery – the two factors separate cleanly as expected.

RIOT appears to have the same problem as the 10-subtest WAIS-IV core battery. Its intended Fluid Reasoning index has only three subtests to begin with (Matrix Reasoning, Visual Puzzles, Figure Weights), and on other professional tests, Visual Puzzles consistently loads almost entirely on Visual-Spatial ability rather than Fluid Reasoning (Wechsler, 2008, 2014; Wechsler et al., 2024). This leaves RIOT's Fluid Reasoning with effectively two unique indicators (Matrix Reasoning, Figure Weights), which is a thin basis for identifying a distinct factor, and weak item-level construction within those subtests could compound the problem further.

Working Memory has an analogous but much larger problem. Visual working memory measures are consistently found in the broader psychometric literature to load primarily on Visual-Spatial ability and only secondarily on Working Memory itself. In the WAIS-V technical manual, for instance, Symbol Span loads .50 on VSI and just .21 on WMI, and Spatial Addition loads .49 on VSI and .26 on WMI, despite both being intended as Working Memory indicators (Wechsler et al., 2024).

A plausible reason is strategic: examinees confronted with a visual working memory task can often visualize the spatial relationships among the items they need to remember rather than holding them in memory the way the task intends, a strategy that is not available on auditory working memory tasks. This may be why auditory subtests like Digit Span are consistently found to be much purer measures of Working Memory. Visual working memory subtests, by contrast, end up measuring visualization strategy – and by extension Visual-Spatial ability – more than they measure working memory itself.

This same pattern shows up in RIOT's own data, especially for Visual Reversal. In the four-factor Schmid-Leiman solution (Table 8), VR loads .38 on the merged Fluid/Spatial factor and shows no particular affinity with Computation Span or Exposure Memory. VR's smaller but still meaningful loading on Processing Speed (.20) also looks like an artifact of administration: the subtest briefly flashes tiles flipping on a grid, one second per tile. The rapid visual scanning this requires likely taps CHC Processing Speed rather than the working memory the subtest is meant to capture.

The other two Working Memory subtests do not fare much better. Exposure Memory is an outright poor indicator of general ability – its communality in the six-factor solution ($h^2 = .21$; Table 9) is the lowest of any subtest – and it correlates only weakly with the rest of the battery ($r = .14$–$.39$; Table 6), including its fellow Working Memory subtests, with which it correlates no more strongly (Computation Span $r = .20$, Visual Reversal $r = .24$) than with subtests from unrelated indices.

Computation Span may have the same problem as WAIS-style Arithmetic subtests. The

two are not identical, but they share both a significant quantitative component and a working-memory component and are conceptually similar in several respects. On the WAIS-V, Arithmetic loads primarily on Fluid Reasoning (.44) and only secondarily on Working Memory (.37) (Wechsler et al., 2024). It is therefore likely that Computation Span splits in a similar way.

In short, none of the three intended working memory subtests appears to primarily measure working memory: VR is a stronger measure of SAI/VSI and PSI than of WMI, CS likely loads more on Gq/quantitative reasoning than on WMI, and EM does not load well on much at all. With no clean indicators left, Working Memory fails to cohere as a distinct factor in any of the factor solutions.

Taken together, these two threads point to the same conclusion from different directions. RIOT's six-factor architecture over-splits the data twice: once by separating Fluid Reasoning from Spatial Ability with too few clean indicators to support the split, and once by carving out a Working Memory factor whose three subtests each measure something other than working memory. Neither problem indicts the CHC framework itself, which does support six distinct broad abilities in principle; the issue lies in RIOT's specific choice and construction of subtests. There are far too many SAI/VSI subtests, not enough FRI subtests, and WMI should be redone from scratch. While a four-factor solution is not what would be theoretically expected within the CHC framework, it is more cohesive than a six-factor model that forces FRI and SAI to split and creates a nonexistent WMI factor.

8 Confirmatory Factor Analysis

8.1 Model Reproduction

The RIOT team published a hierarchical six-factor model, with the six index factors loading onto a higher-order general intelligence factor (g), as RIOT's underlying structure (Warne, 2025a). This section uses that model as the baseline against which to formally test the conclusions drawn from the Intercorrelations and EFA sections above – namely, whether the four-factor structure favored by parallel analysis, BIC, and inspection of the heatmap actually outperforms RIOT's own six-factor architecture once both are fit in a confirmatory framework.

Before comparing this structure to alternative models, we first attempted to reproduce it directly. Alongside global fit statistics, the bulletin presents a path diagram with each subtest's standardized loading on its index and each index's standardized loading on g (Warne, 2025a), allowing us to check the reproducibility of both model fit and individual parameter estimates.

We specified the model exactly as described in the bulletin: each of the six indices (Verbal Reasoning, Fluid Reasoning, Spatial Ability, Working Memory, Processing Speed, Reaction Time) as a first-order latent factor with its constituent subtests as indicators (per Table 2), and a second-order general factor (g) on which all six index factors load. The model was fit in `lavaan` using maximum likelihood estimation on the subtest intercorrelation matrix (Table 6), with $N = 417$ matching the norm sample size reported in the bulletin. Because only the intercorrelation matrix is available, multivariate normality could not be tested or corrected for, so all CFA results assume multivariate normality.

Table 10 compares the fit statistics reported in the bulletin to our reproduction.

Table 10: Fit statistics: Reported vs. reproduced hierarchical six-factor model

	Reported (Warne, 2025a)	Reproduced
χ^2	234.36	238.97
df	98	98
CFI	.950	.949
RMSEA	.058	.059
RMSEA 90% CI	[.048, .067]	[.049, .068]
SRMR	.045	.048

Note. Model fit in `lavaan` (ML estimation) to the intercorrelation matrix in Table 6, $N = 417$.

Our reproduction matches the reported fit statistics closely: the χ^2 difference is small (238.97 vs. 234.36 on the same 98 degrees of freedom) and likely attributable to rounding, and CFI, RMSEA, and SRMR all agree with the bulletin's values to within 0.003.

Table 11 compares the standardized loadings depicted in the bulletin's path diagram to those obtained in our reproduced model.

Table 11: Standardized loadings: Reported vs. reproduced CFA loadings

Subtest/Factor	Path	Reported loading	Reproduced loading	Difference
Vocabulary	VO → VRI	0.79	0.79	0.00
Information	IN → VRI	0.77	0.77	0.00
Analogies	AN → VRI	0.73	0.73	0.00
Matrix Reasoning	MR → FRI	0.76	0.77	+0.01
Visual Puzzles	VP → FRI	0.72	0.72	0.00
Figure Weights	FW → FRI	0.78	0.78	0.00
Object Rotation	OR → SAI	0.70	0.70	0.00
SToVeS	STV → SAI	0.72	0.72	0.00
Spatial Orientation	SO → SAI	0.85	0.85	0.00
Computation Span	CS → WMI	0.50	0.50	0.00
Exposure Memory	EM → WMI	0.44	0.44	0.00
Visual Reversal	VR → WMI	0.59	0.59	0.00
Symbol Search	SS → PSI	0.85	0.85	0.00
Abstract Matching	AM → PSI	0.57	0.57	0.00
Simple Reaction Time	SRT → RTI	0.86	0.87	+0.01
Choice Reaction Time	CRT → RTI	0.84	0.84	0.00
Verbal Reasoning	VRI → g	0.76	0.76	0.00
Fluid Reasoning	FRI → g	0.92	0.92	0.00
Spatial Ability	SAI → g	0.98	0.98	0.00
Working Memory	WMI → g	0.92	0.93	+0.01
Processing Speed	PSI → g	0.58	0.58	0.00
Reaction Time	RTI → g	0.51	0.51	0.00

Note. Reported loadings are read from the standardized path diagram in Warne (2025a). Reproduced loadings are standardized loadings from our CFA.

The two sets of loadings agree almost exactly across both levels of the model: 19 of 22 paths match the reported value to two decimal places, and the three that differ are within rounding error.

Figure 6 depicts the full path diagram of our reproduced model with standardized loadings and residual variances.

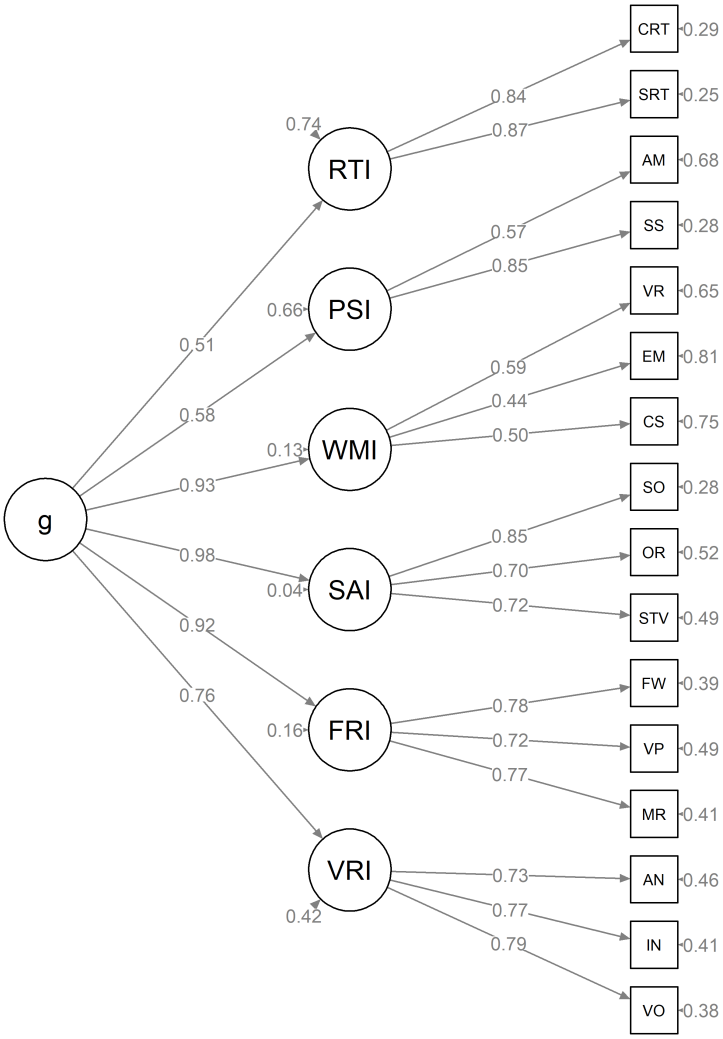


Figure 6: Path diagram of the reproduced hierarchical six-factor null model

For transparency and reproducibility, the complete `lavaan` model output – including unstandardized estimates, standard errors, z-values, and all fit indices – is reported in Appendix C.

Overall, both the global fit statistics and the individual standardized loadings – at the subtest level and the index level – closely reproduce the values reported in the developer's bulletin. We therefore conclude that we have successfully reproduced RIOT's hierarchical six-factor model from the recovered intercorrelation matrix.

8.2 Competing Models

Having reproduced the developer's hierarchical six-factor model in Section 8.1, we use our reproduced version of that model as the null model against which alternative factor structures (e.g., a single-factor model and alternative hierarchical groupings) are compared below.

Model 1: Single-factor model. As a baseline point of comparison, we fit a single-factor model in which all 16 subtests load directly onto one common factor, with no intermediate index-level structure. The model was fit in `lavaan` (ML estimation) on the same subtest intercorrelation matrix (Table 6), with $N = 417$. This model fit the data poorly relative to the reproduced hierarchical six-factor model: $\chi^2(104) = 712.91$, $p < .001$, CFI = .778, TLI = .743, RMSEA = .118 (90% CI [.110, .127]), SRMR = .078. All standardized loadings on the single factor were positive and significant ($p < .001$), ranging from .35 (Abstract Matching) to .83 (Spatial Orientation). Figure 7 shows the corresponding path diagram.

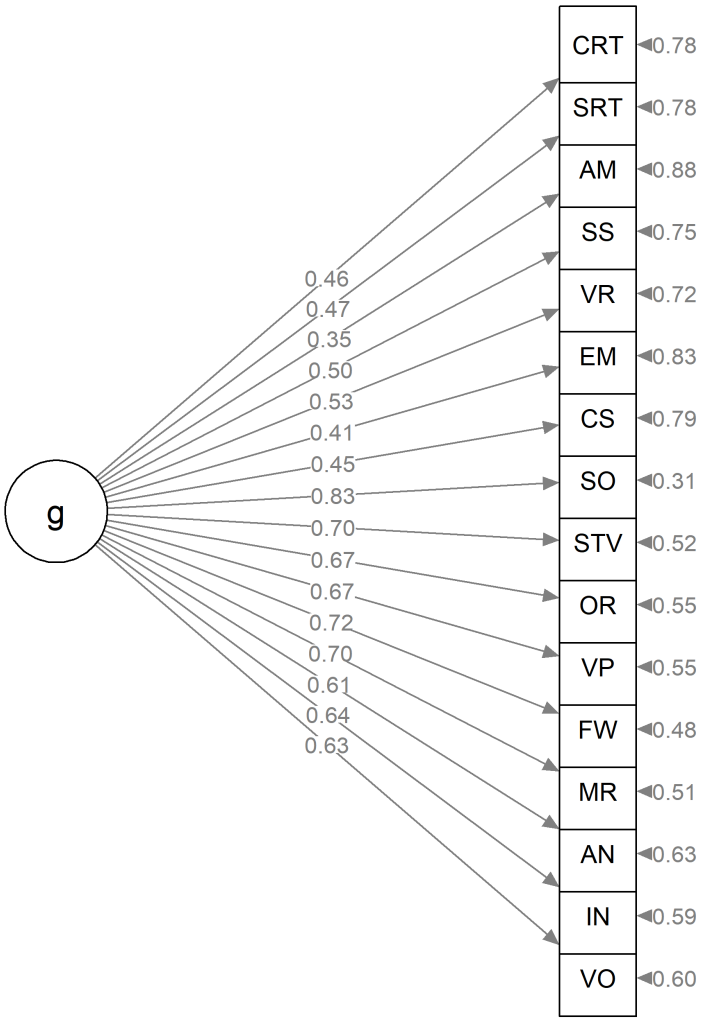


Figure 7: Path diagram of Model 1 (single-factor)

Model 2: Four-factor hierarchical model (EFA-informed). Motivated by the four-factor solution preferred by both parallel analysis and BIC in the EFA (Section 7), we fit a second hierarchical model that keeps the second-order g factor but replaces the six a priori indices with the four EFA-derived group factors: Verbal (VO, IN, AN), Reaction Time (SRT, CRT), Processing Speed (SS, AM), and a single merged Perceptual Reasoning factor (PRI)

combining all nine Fluid Reasoning, Spatial Ability, and Working Memory subtests (MR, FW, OR, VP, STV, VR, SO, EM, CS). Spatial Orientation was additionally allowed to cross-load onto Verbal Reasoning, consistent with the subtest's content and with the cross-loadings found in both the four- and six-factor EFAs. Folding the "working memory" subtests into PRI is not ideal, but the EFA results suggest that they do not form a coherent factor of their own, and the merged PRI factor is the most reasonable place for them.

The model was fit in `lavaan` (ML estimation, `std.lv = TRUE`) on the same intercorrelation matrix, $N = 417$, with 37 free parameters. Fit was good: $\chi^2(99) = 192.58$, $p < .001$, CFI = .966, TLI = .959, RMSEA = .048 (90% CI [.038, .058]), SRMR = .037. Figure 8 shows the corresponding path diagram.

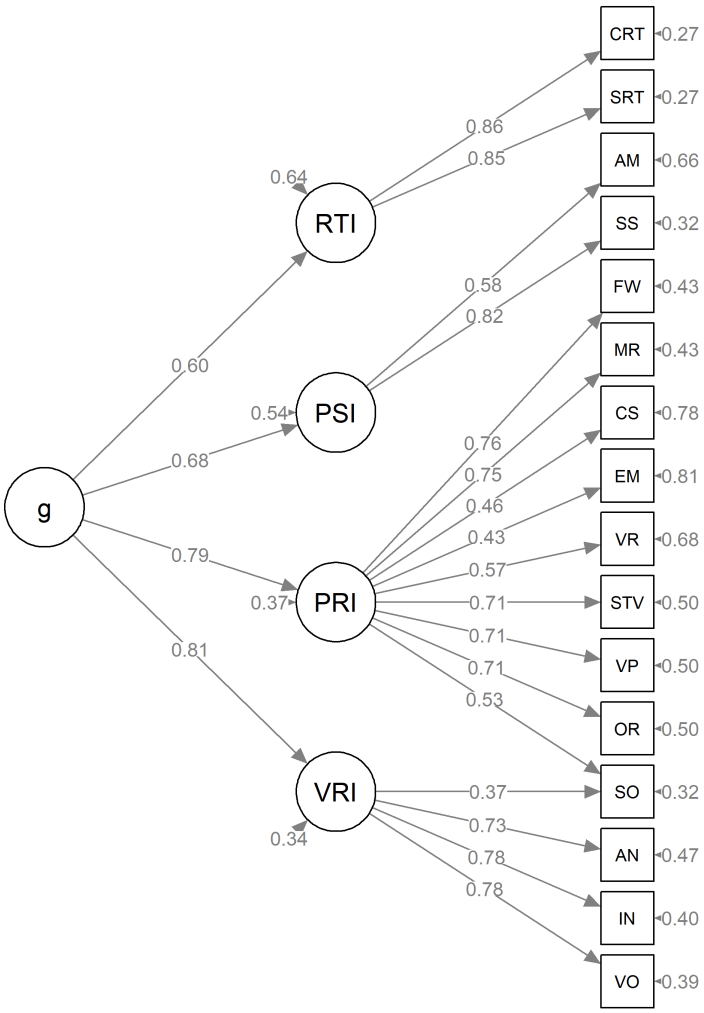


Figure 8: Path diagram of Model 2 (EFA-informed hierarchical four-factor)

Model 2 is not nested within the null model – it groups subtests differently rather than merely constraining the null model's parameters – so a chi-square difference test does not apply; the two can instead be compared on fit indices and information criteria. Model 2 outperforms the null model on every index reported here (CFI .966 vs. .949; RMSEA .048 vs. .059; SRMR .037 vs. .048; Table 12) while using one fewer free parameter (37 vs. 38), and both information criteria prefer it: AIC = 16326.1 vs. 16374.5, and BIC = 16475.3 vs. 16527.7. However, because Model 2 was derived from an EFA of this same matrix, this comparison is in-sample and is treated as supporting evidence (Section 8.3).

Model 3: Five-factor hierarchical model (CHC-based split of Fluid/Spatial). Model 2's merged Perceptual Reasoning factor collapses Fluid Reasoning and Spatial Ability together, which CHC theory treats as distinct broad abilities (Gf and Gv, respectively; Table 1). To test whether the two can be statistically separated, we fit a third model that splits PRI into a Fluid Reasoning factor and a Visual-Spatial factor. Spatial Orientation loads on Visual-Spatial and keeps its cross-loading onto Verbal Reasoning, for the reasons described under Model 2. FRI comprises three subtests: MR, FW, and CS. Computation Span was assigned to Fluid Reasoning because it is plausibly a primarily quantitative (Gq) reasoning task, which CHC theory often places within Gf.

The subtests assigned to VSI were OR, VP, STV, VR, EM, and SO. Visual Puzzles was moved from FRI (its placement in the null model) to VSI because, on other professional intelligence tests, it consistently loads more strongly on Visual-Spatial ability than on Fluid Reasoning (Wechsler, 2008, 2014; Wechsler et al., 2024). Visual Reversal was also placed under VSI, for reasons discussed in the EFA Discussion (Section 7.4). Lastly, Exposure Memory was placed under VSI because it is, by description, a visual working memory task: recognizing which images were shown in a briefly presented sequence (Table 2).

The model was fit in `lavaan` (ML estimation, `std.lv = TRUE`) on the same intercorrelation matrix, $N = 417$, with 38 free parameters – coincidentally the same count as the null model, since one added cross-loading offsets one fewer second-order path. Fit was good and better than the null model at identical model complexity on every index reported here except SRMR, where the two tie: $\chi^2(98) = 219.59$, $p < .001$, CFI = .956, TLI = .946, RMSEA = .055 (90% CI [.045, .064]), SRMR = .048, AIC = 16355.1, BIC = 16508.3 (vs. the null model's $\chi^2(98) = 238.97$, CFI = .949, RMSEA = .059, SRMR = .048, AIC = 16374.5, BIC = 16527.7). Figure 9 shows the corresponding path diagram.

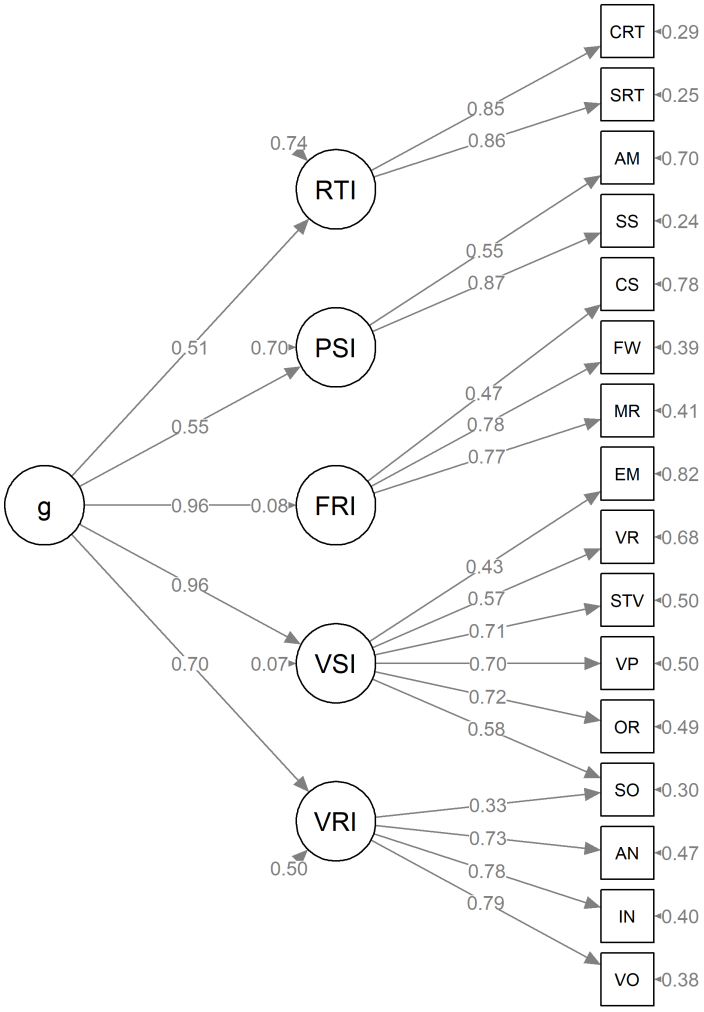


Figure 9: Path diagram of Model 3 (CHC-based hierarchical five-factor)

Table 12: Fit indices for competing CFA models

Model	χ^2	df	CFI	RMSEA	RMSEA 90% CI	SRMR
Null model (reproduced hierarchical six-factor)	238.97	98	.949	.059	[.049, .068]	.048
Model 1 (single-factor)	712.91	104	.778	.118	[.110, .127]	.078
Model 2 (EFA-informed hierarchical four-factor)	192.58	99	.966	.048	[.038, .058]	.037
Model 3 (CHC-based hierarchical five-factor)	219.59	98	.956	.055	[.045, .064]	.048

Table 12 lines up all four models directly. Model 1 sits far below the other three on every index (CFI .778, RMSEA .118); the remaining three cluster together in a materially better range. Among those three, Model 2 posts the best fit on every index reported (CFI .966, RMSEA .048, SRMR .037), Model 3 falls in between (CFI .956, RMSEA .055, SRMR .048 – tying the null model's SRMR exactly), and the null model brings up the rear (CFI .949, RMSEA .059) despite still clearing conventional adequate-fit thresholds.

8.3 CFA Discussion

The CFA results reinforce the exploratory findings. RIOT's published hierarchical six-factor model reproduces cleanly, so the relevant question is not whether that structure can be reproduced, but whether its six first-order factors are empirically distinct from one another. The strongest evidence on that question comes from RIOT's own model rather than from any alternative to it.

In the reproduced six-factor model (Figure 6), Fluid Reasoning, Spatial Ability, and Working Memory load .92, .98, and .93 on g. To estimate how strongly these three factors correlate without relying on the second-order structure, we also fit the same six first-order factors as freely correlated factors, with no g. This gave correlations of .90 (95% CI [.86, .95]) between FRI and SAI, .93 ([.84, 1.03]) between FRI and WMI, and .88 ([.78, .98]) between SAI and WMI. Under the guidelines of Rönkkö and Cho (2022), which classify factor correlations by the upper limit of their 95% CI, the FRI–SAI and SAI–WMI correlations indicate a moderate discriminant validity problem (upper limits between .9 and 1), and the FRI–WMI correlation a severe one (upper limit ≥ 1), meaning FRI and WMI cannot be distinguished empirically. Factors this close to g and to one another provide little basis for interpreting FRI, SAI, and WMI as separate index scores. Verbal Reasoning, Processing Speed, and Reaction Time, by contrast, load .76, .58, and .51 on g.

Merging the three into a single Perceptual Reasoning factor (Model 2; Figure 8) resolves this: it loads .79 on g, comparable to Verbal Reasoning (.81). This model also fits better than the null model on every reported index with one fewer free parameter (Table 12). Because it was specified from an EFA of the same matrix, some of this advantage is expected, and confirming it would require an independent sample.

The single-factor model's collapse (CFI = .778) is informative in the other direction. RIOT clearly contains a strong general factor, consistent with the large first eigenvalue in the EFA, but g alone cannot explain the subtest correlations. The real question, then, is not whether RIOT measures g, but how many narrower abilities can be defensibly distinguished underneath it.

Model 3's primary aim was to test whether Fluid Reasoning and Visual-Spatial ability can be statistically separated in line with CHC theory, which treats Gf and Gv as distinct broad abilities. Even after Model 2's PRI subtests were reassigned to FRI or VSI based on the EFA results and their contents, the separated model still fit worse than the merged model, mirroring the EFA's own failure to recover a clean Fluid/Spatial split.

As discussed in the EFA Discussion (Section 7.4), this is not evidence that Gf and Gv are indistinguishable abilities in general. More likely, RIOT's present subtests do not supply sufficiently distinct indicators of the two constructs, a problem that weaknesses in individual item design may compound. The merged Perceptual Reasoning factor should therefore be read as a statistical description of how RIOT's present subtests behave rather than as a new CHC broad ability. The FRI and SAI indicators share largely homogeneous covariance with one another, and placing the Working Memory subtests within PRI is the best (if imperfect) way to handle their failure to form a coherent factor of their own.

Model 3 is nonetheless informative in its own right: relocating CS, EM, and VR out of Working Memory improves fit even at identical model complexity to the null model. This is consistent with three other observations: the three subtests correlate only .20–.32 with one another (Table 6), only Computation Span retains a Working Memory loading in the EFA, and the WMI factor's correlation with FRI cannot be distinguished from 1.00. Together, these observations indicate that CS, EM, and VR do not form a distinct Working Memory factor.

Further evidence of the same low discriminant validity among FRI, SAI, and WMI shows up on the other side of the model. The other three factors – VRI, PSI, and RTI – all become *more* g-loaded once FRI, SAI, and WMI stop being modeled as separate factors. Their composite g-loadings rise from .683 to .728 (VRI), .479 to .553 (PSI), and .466 to .552 (RTI). This is

a purely second-order effect: their subtest → factor (first-order) loadings barely move, while their factor → g (second-order) loadings rise, indicating that the entire gain comes from how the second-order structure repartitions shared variance. This could happen only if the null model's forced FRI/SAI/WMI split was distorting not just those three factors, but how the whole model allocates variance to g. Once shared variance between the FRI, SAI, and WMI subtests is no longer being forced through g, the other factors are able to better approximate their true relationship to g, removing the previous artificial suppression.

Spatial Orientation is the clearest individual casualty of this distortion. Because the null model gives SO no direct path to Verbal Reasoning, its strong correlation with the Verbal subtests can only be explained by routing that covariance through g – inflating SAI's second-order loading and SO's own indirect loading on g to .83 (Table 11). Model 2 instead lets SO cross-load directly onto Verbal Reasoning, modeling that shared variance explicitly rather than routing it through g. As a result, its loading on g falls to a more realistic .72 (Table 14) – still the highest of any subtest, but no longer inflated.

A further consequence of this over-partitioning is that it inflates RIOT's own implied estimate of how strongly g saturates the FSIQ. Because FRI, SAI, and WMI are forced to remain separate factors despite their high mutual correlations, their shared covariance is routed through g instead, inflating the estimated FSIQ loading on g to .92, compared to Model 2's more modest .86. Given the evidence above that FRI, SAI, and WMI are not well differentiated, Model 2's .86 is the better-supported estimate.

Taken together, these results tell a consistent story. RIOT measures a strong general factor. Beyond it, only some of RIOT's six proposed indices hold up as genuinely separable dimensions: Verbal Reasoning, Processing Speed, and Reaction Time are well-supported, but Fluid Reasoning, Spatial Ability, and Working Memory are not. FRI and SAI resist separation and instead behave as a single domain, while WMI does not cohere to begin with. Forcing that domain into three factors anyway leaves each of them nearly indistinguishable from g and distorts every g-loading in the model, inflating FRI's, SAI's, and WMI's own loadings (and, in turn, the FSIQ's) while artificially suppressing VRI, PSI, and RTI. Model 2's four-factor structure avoids all of this, and also fits these data better with fewer parameters – though, as an in-sample comparison, that fit advantage is supporting rather than primary evidence.

9 Subtest and Composite Loadings

Section 8.3 concluded that Model 2 – the four-factor hierarchical CFA that replaces RIOT's six a priori indices with Verbal Reasoning (VRI), Reaction Time (RTI), Processing Speed (PSI), and a single merged Perceptual Reasoning factor (PRI) combining the Fluid Reasoning, Spatial Ability, and Working Memory subtests – is the best-supported structure for RIOT's data. Table 13 reports each composite's g-loading under this model, and Table 14 reports each subtest's g-loading, grouped by the Model 2 factor it belongs to.

Table 13: Composite loadings on g (Model 2)

Index/Score	Loading on g
Full-Scale IQ (FSIQ)	0.86
Verbal Reasoning (VRI)	0.73
Perceptual Reasoning (PRI)	0.76
Processing Speed (PSI)	0.55
Reaction Time (RTI)	0.55

Note. For reference, if Model 2's subtest g-loadings are used to compute composites under RIOT's original six-index definitions, the Fluid Reasoning, Spatial Ability, and Working Memory composites load 0.69, 0.73, and 0.54 on g, respectively.

Table 14: Subtest loadings on g (Model 2)

Subtest	Factor	Loading on g
Vocabulary	VRI	0.63
Information	VRI	0.63
Analogies	VRI	0.59
Matrix Reasoning	PRI	0.59
Visual Puzzles	PRI	0.56
Figure Weights	PRI	0.60
Object Rotation	PRI	0.56
SToVeS	PRI	0.56
Spatial Orientation	PRI	0.72
Computation Span	PRI	0.36
Exposure Memory	PRI	0.34
Visual Reversal	PRI	0.45
Symbol Search	PSI	0.56
Abstract Matching	PSI	0.39
Simple Reaction Time	RTI	0.51
Choice Reaction Time	RTI	0.52

The Recommendations section that follows adopts Model 2 as the basis for its guidance, along with the index- and subtest-level g-loadings reported in Tables 13 and 14 above.

10 Recommendations

Unlike a standard clinical IQ test, which is published once and revised only at the next edition, RIOT is administered entirely online and can in principle be revised at any time. The recommendations below translate the findings of this report into concrete priorities for the RIOT team.

These recommendations are not exhaustive; they are grounded only in the summary-level statistics available. Item-level analysis – item difficulty, discrimination, and so on – would very plausibly surface additional problems, and could well suggest more specific solutions than the ones proposed here for the same subtests. With that caveat, the evidence assembled in this report supports the following priorities.

Verbal Reasoning. VRI is the healthiest of RIOT's six proposed indices and does not need structural change. Its three subtests hang together cleanly in both the EFA (Table 9) and CFA (Table 11), and under the best-supported Model 2 structure, the VRI factor carries a solid .81 loading on g (Figure 8), in line with Gc factors in the psychometric literature. There is room for incremental item-quality improvements – Vocabulary, Information, and Analogies each load only .59–.63 on g individually (Table 14), leaving some headroom relative to a well-built Gc subtest – but these refinements are a lower priority than the issues below.

Fluid Reasoning and Spatial Ability. Matrix Reasoning, Figure Weights, and Visual Puzzles all load on g well below where their direct equivalents land on the WAIS-V and WISC-V, as Table 15 shows.

Table 15: Subtest loadings on g: WAIS-V vs. WISC-V vs. RIOT

Subtest	WAIS-V	WISC-V	RIOT
Figure Weights	0.78	0.67	0.60
Matrix Reasoning	0.73	0.67	0.59
Visual Puzzles	0.74	0.69	0.56

Note. WAIS-V loadings (Wechsler et al., 2024); WISC-V loadings (Wechsler, 2014); RIOT loadings (Table 14).

This is consistent with the rest of this report's findings on RIOT's Fluid Reasoning and Spatial Ability subtests (Sections 7.4 and 8.3): as currently built, none of the three comes close to isolating a broad ability as effectively as its professional-test counterparts do. Because item-level data were not available, this report cannot specify exactly what item-construction changes would close that gap. What it can say is where each subtest should be aimed. Matrix Reasoning and Figure Weights should be revised to more purely isolate Gf; Visual Puzzles should be revised to more purely isolate Gv.

Visual Puzzles should also be moved out of Fluid Reasoning and into Spatial Ability, a move on which CHC theory and the professional-test literature converge. Visual Puzzles' WAIS-V and WISC-V equivalents are both classified as core Visual-Spatial subtests, not Fluid Reasoning subtests (Wechsler, 2014; Wechsler et al., 2024), and Weiss et al. (2013) found the same pattern in their five-factor re-analysis of the WAIS-IV. EFA of RIOT's own data agrees: the six-factor EFA, which models Spatial Ability as its own factor, shows Visual Puzzles cross-loading onto Spatial Ability nearly as strongly as it loads onto Fluid Reasoning. Leaving it under Fluid Reasoning also leaves FRI with effectively only two unique indicators, Matrix Reasoning and Figure Weights, which is exactly the thin-indicator problem that kept Gf and Gv from separating on the original 10-subtest WAIS-IV core battery (Section 7.4). Model 3's underperformance relative to Model 2 (Table 12) suggests the same problem is at work in RIOT.

RIOT has far too many Spatial Ability subtests relative to Fluid Reasoning subtests. SAI already has three (Spatial Orientation, SToVeS, and Object Rotation); Visual Puzzles makes four, and Visual Reversal – which behaves mainly as a Gv measure – would make five subtests all primarily measuring Gv, while Fluid Reasoning is left with only two. Simply adding new Fluid Reasoning subtests on top of the existing battery would fix that imbalance but at the cost of a longer test, working against the administration time RIOT is aiming for. A better path would be to introduce well-constructed Fluid Reasoning subtests to round out FRI while retiring one or two of Spatial Ability's five subtests to make room for them.

Working Memory. The current Working Memory index requires substantial redevelopment. None of its three subtests provides a sufficient measure of working memory, and none is a particularly strong measure of g: Computation Span, Exposure Memory, and Visual Reversal load just .36, .34, and .45 on g respectively under Model 2 (Table 14) – among the lowest loadings in the entire battery. These findings suggest that the current subtests are good measures of neither working memory nor general ability.

The underlying cause, discussed at length in Section 7.4, is that all three current subtests are visual/reasoning tasks, and visual working-memory tasks are consistently found in the literature to load more heavily on Visual-Spatial ability than on Working Memory itself – often roughly twice as much, as with the WAIS-V's Symbol Span (.50 VSI vs. .21 WMI) and Spatial Addition (.49 VSI vs. .26 WMI) (Wechsler et al., 2024). One fix, then, would be to replace the current visual Working Memory subtests with auditory ones. Auditory working-memory tasks such as Digit Span are consistently found to be the purest measures of Gwm, likely because they do not give examinees a visualization strategy to lean on the way a visual grid or image sequence does.

Processing Speed. Symbol Search is a well-established subtest type in the psychometric literature and performs accordingly here, loading cleanly in both the EFA (Table 9) and CFA. Abstract Matching, however, does not: it loads only .39 on g and .58 on its own PSI factor under Model 2 – a weak loading on the very factor it is supposed to measure. This report therefore recommends that Abstract Matching be scrapped and replaced with a better-established measure of processing speed.

Reaction Time. Simple and Choice Reaction Time load nearly identically on g under Model 2 (.51 and .52 respectively; Table 14). Choice reaction time is generally more g-loaded than simple reaction time, because choosing between stimuli adds a decision component, but this increase depends on the number of choices. Across 15 samples tested on the Hick paradigm, the correlation between reaction time and IQ increased roughly linearly from two to four to eight choices (Jensen, 2006, pp. 164–165). Two-choice reaction time is the closest to simple reaction time, adding only a single increment of g-loaded decision processing. In a sample spanning a wide range of ability (IQs 57 to 130), the correlation with IQ moved only from roughly $-.50$ for simple reaction time to $-.56$ for two-choice, whereas four-, six-, and eight-choice reaction time reached roughly $-.65$ to $-.75$ (negative because faster is better; Lally & Nettelbeck, 1977, as reported in Jensen (2006, p. 165)). In other words, simple and two-choice reaction time measure nearly the same thing, because the decision component that two-choice RT adds is small. RIOT's data show the same pattern: Simple and Choice Reaction Time correlate very strongly (Table 6), and together with their nearly identical g-loadings, this indicates that RIOT's two-choice task adds little beyond Simple Reaction Time. This report therefore recommends changing Choice Reaction Time from a two-choice to a four-choice format, which would better distinguish it from Simple Reaction Time, make it a stronger measure of g, and isolate decision time more cleanly.

Another possible concern is that the very strong correlation between the two ($r = .73$; Table 6) may partly reflect a feature unique to online testing. Examinees complete both tasks under the same technical environment, so hardware-specific differences – display, browser, operating-system, and input-device effects – could contribute correlated measurement error to both scores. This can be represented as a shared, person-specific hardware-latency term H:

$$SRT_{\text{obs}} = SRT_{\text{ability}} + H + e_S, \qquad CRT_{\text{obs}} = CRT_{\text{ability}} + H + e_C.$$

Assuming, reasonably, that H is independent of reaction-time ability and of the error terms, we obtain

$$\text{Cov}(SRT, CRT) = \text{Cov}(SRT_{\text{ability}}, CRT_{\text{ability}}) + \text{Var}(H),$$

so between-examinee variance in hardware latency mechanically inflates the observed SRT–CRT covariance on top of whatever covariance genuine reaction-speed ability contributes. Because Reaction Time has only two subtests, any systematic variance they share, whether cognitive or technical, is absorbed into the RTI factor; with only two indicators, there is no way to separate shared reaction-speed variance from shared technical variance or to estimate how large this effect is.

Subtest-Level g Loadings. Figure 10 places all of these subtest-level findings side by side against WAIS-V and WISC-V. RIOT's subtest g-loadings cluster heavily in the .50–.59 band, which alone accounts for half of its 16 subscores, while both professional tests are centered a full band higher, at .60–.69 (45% of WAIS-V subtests, 50% of WISC-V subtests). The gap is starkest at the top end of the distribution: 30% of WAIS-V subtests and 19% of WISC-V subtests reach the .70–.79 band, versus just 6% of RIOT's. RIOT also has a larger share of subtests in the weak .30–.39 band than either professional test (19%, versus 5% for WAIS-V and 6% for WISC-V).

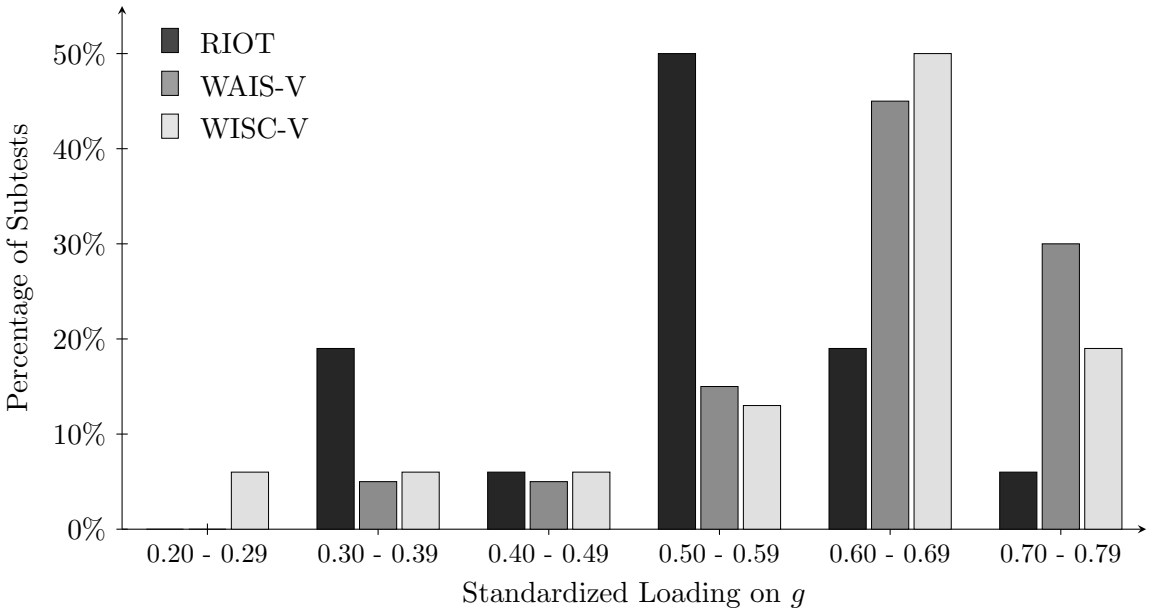


Figure 10: Distribution of subtest loadings on g: RIOT vs. WAIS-V vs. WISC-V

These subtest-level differences carry through to how well each test measures g overall: RIOT's FSIQ g-loading is .86, while the WAIS-V's and WISC-V's are .94 and .92, respectively. RIOT is a decent measure of g, but if it wants to stand alongside the gold-standard professional IQ tests it claims as peers, it needs to improve its subtests so that they are more efficient indicators of g.

Composite Norms. The artificial ceilings and floors that RIOT imposes on its composites should be substantially widened or dropped altogether. As discussed in Section 3.3, clipping the FSIQ and indices to the same range as individual subtests ignores the composite effect and unnecessarily discards valuable information at both tails. Removing the bounds, or at minimum widening them to reflect the true range of composite scores, would allow more accurate scores for examinees at both ends of the ability distribution.

Transparency. RIOT was launched with an explicit pledge to be "the most transparent test ever created" (Warne, 2025c) and "completely trustworthy and auditable" (Warne, 2025e). More than a year after launch, that pledge remains unfulfilled: other than a single bulletin, no data or analysis code has been released, and the manual sits behind a request-access process. The single most valuable change the RIOT team could make, independent of any subtest fix above, is to make good on that original promise.

11 Conclusion

This report finds that RIOT measures a strong general cognitive factor. However, the evidence does not support all six reported indices equally well. Across both exploratory and confirmatory analyses, the data favor a four-factor structure: in RIOT's own six-factor model, Fluid Reasoning, Spatial Ability, and Working Memory are nearly indistinguishable from g and from one another, while Verbal Reasoning, Processing Speed, and Reaction Time emerge as the clearest distinct domains.

The main structural weaknesses involve Fluid Reasoning, Spatial Ability, and especially Working Memory. FRI and SAI do not separate cleanly in RIOT's present battery, while the three WMI subtests fail to form a coherent Working Memory factor at all. This directly calls into question the construct-level interpretability of the FRI and SAI composites, and particularly the WMI composite, as distinct measures of the abilities indicated by their labels. The issue is not simply whether these scores can be calculated reliably, but whether differences in them can defensibly be interpreted as genuinely reflecting distinct Fluid Reasoning, Spatial Ability, and Working Memory abilities. The evidence assembled here supports considerably greater confidence in RIOT's FSIQ, VRI, PSI, and RTI than in these three index scores.

Ultimately, this report aims to improve transparency regarding RIOT, particularly with respect to its methodology, norming, scoring, and internal structure. While its tone has at times been critical, that criticism is intended to be constructive rather than dismissive. RIOT is an ambitious project, and the recommendations offered throughout this report are meant not only to identify areas where the present evidence falls short, but also to suggest ways the instrument might be strengthened. More broadly, it is hoped that the findings presented here will prove useful as RIOT continues to be evaluated, refined, and improved.

References

de Leeuw, J., & Mair, P. (2009). Multidimensional scaling using majorization: SMACOF in R. *Journal of Statistical Software*, *31*(3), 1–30. https://doi.org/10.18637/jss.v031.i03

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185.

Jensen, A. R. (2006). *Clocking the mind: Mental chronometry and individual differences*. Elsevier.

Kruskal, J. B. (1964). Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, *29*(1), 1–27. https://doi.org/10.1007/BF02289565

Marshalek, B., Lohman, D. F., & Snow, R. E. (1983). The complexity continuum in the radex and hierarchical models of intelligence. *Intelligence*, *7*(2), 107–127. https://doi.org/10.1016/0160-2896(83)90023-5

Mosier, C. I. (1943). On the reliability of a weighted composite. *Psychometrika*, *8*(3), 161–168.

Revelle, W. (2025). psych: Procedures for psychological, psychometric, and personality research [R package version 2.5]. https://CRAN.R-project.org/package=psych

Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational Research Methods*, *25*(1), 6–14. https://doi.org/10.1177/1094428120968614

Schmid, J., & Leiman, J. M. (1957). The development of hierarchical factor solutions. *Psychometrika*, *22*(1), 53–61.

Warne, R. T. (n.d.). *The RIOT IQ test manual*. Riot IQ. Retrieved August 15, 2026, from https://www.riotiq.com/test-manual

Warne, R. T. (2025a). *Informational Bulletin for the Reasoning and Intelligence Online Test* (RIOT Technical Report No. 2025-01). RIOT IQ.

Warne, R. T. (2025b). Reasoning and intelligence online test, version 1.0. *RIOT IQ*. https://riotiq.com

Warne, R. T. (2025c, February). *What is the RIOT IQ test?* Riot IQ. https://www.riotiq.com/articles/riot-specific-information/what-is-the-riot

Warne, R. T. (2025d, February 20). *The 15 subtests of the RIOT*. RIOT IQ. https://www.riotiq.com/articles/riot-specific-information/the-15-subtests-of-riot

Warne, R. T. (2025e, March 20). *Comment on: Can anyone recommend valid and reliable online IQ tests?* ResearchGate. https://www.researchgate.net/post/Can-anyone-recommend-valid-and-reliable-online-IQ-tests

Warne, R. T. (2025f, July 18). *Available now: The reasoning and intelligence online test*. Riot IQ. https://www.riotiq.com/articles/riot-specific-information/available-now-the-reasoning-and-intelligence-online-test

Warne, R. T. (2025g, December 3). *Are online IQ tests valid?* Riot IQ. https://www.riotiq.com/articles/online-iq-tests/are-online-iq-tests-valid

Warne, R. T. (2026). Heterogeneous relationships between working speed and ability on the reasoning and intelligence online test (RIOT). *Intelligence*, *115*, 101990. https://doi.org/10.1016/j.intell.2025.101990

Wechsler, D. (2008). *Wechsler adult intelligence scale: Technical and interpretive manual* (4th ed.). Psychological Corporation.

Wechsler, D. (2014). *Wechsler intelligence scale for children: Technical and interpretive manual* (5th ed.). NCS Pearson.

Wechsler, D., Raiford, S. E., & Presnell, K. (2024). *Wechsler adult intelligence scale: Technical and interpretive manual* (5th ed.). NCS Pearson.

Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013). WAIS-IV and clinical validation of the four- and five-factor interpretative approaches. *Journal of Psychoeducational Assessment*, *31*(2), 94–113. https://doi.org/10.1177/0734282913478030

AI assistance was used to help draft and edit the prose of this report. All analysis, data processing, research, and conclusions are the author's own.

Appendix A Raw-to-T-Score Lookup Tables

Table A.1: Raw-to-T-score conversion, VRI, FRI, and SAI subtests, Age 18–39

Raw	VO	IN	AN	MR	VP	FW	OR	STV	SO
0		31	31					33	31
1		33	32					35	32
2		35	33					37	34
3		37	35			31		40	36
4	31	39	36		31	33		42	38
5	34	41	38	31	33	35		44	39
6	36	43	40	32	35	37	31	46	41
7	38	45	41	34	37	40	34	48	43
8	41	47	43	36	38	42	36	50	45
9	43	49	45	38	40	45	38	52	47
10	45	51	46	40	42	47	40	54	49
11	47	53	48	42	43	50	43	56	51
12	50	55	49	44	45	52	45	58	52
13	52	57	51	46	47	55	47	60	54
14	54	58	53	48	49	57	50	62	56
15	56	60	54	50	50	60	52	64	58
16	59	62	56	52	52	62	54	66	60
17	61	64	58	54	54	65	56	68	62
18	63	66	59	56	55	67	59	70	64
19	65	68	61	58	57	69	61	72	65
20	68	70	62	60	59	72	63	74	67
21	70	72	64	61	61	74	66	77	69
22	72	74	66	63	62	77	68	79	71
23	75	76	67	65	64	79	70	80	73
24	77	78	69	67	66	80	73		75
25	79	80	71	69	68		75		77
26	80		72	71	69		77		78
27			74	73	71		79		80
28			75	75	73		80		
29			77	77	74				
30			79	79	76				
31			80	80	78				
32					80				

Note. AN = Analogies; FW = Figure Weights; IN = Information; MR = Matrix Reasoning; OR = Object Rotation; SO = Spatial Orientation; STV = SToVeS; VP = Visual Puzzles; VO = Vocabulary. Blank cells indicate raw scores beyond RIOT's 31T and 80T hard cutoffs. The actual subtest may contain fewer questions than are needed to reach the 80T ceiling.

Table A.2: Raw-to-T-score conversion, VRI, FRI, and SAI subtests, Age 40–59

Raw	VO	IN	AN	MR	VP	FW	OR	STV	SO
0		31	31					34	31
1		32	32					36	32
2		34	33					38	34
3	31	36	35			31		41	36
4	32	38	36	31	31	33		43	38
5	34	40	38	33	32	35		45	40
6	36	41	39	35	34	38	31	47	41
7	38	43	41	37	36	40	32	50	43
8	40	45	42	39	37	43	35	52	45
9	42	47	44	41	39	45	37	54	47
10	43	49	45	43	41	48	40	56	49
11	45	50	47	45	43	51	43	59	51
12	47	52	48	47	44	53	45	61	53
13	49	54	50	48	46	56	48	63	55
14	51	56	51	50	48	58	51	66	57
15	53	58	53	52	50	61	53	68	59
16	55	59	54	54	52	64	56	70	61
17	57	61	56	56	53	66	59	72	63
18	59	63	57	58	55	69	61	75	65
19	61	65	59	60	57	71	64	77	67
20	63	67	60	62	59	74	67	79	69
21	65	69	62	64	60	76	69	80	71
22	67	70	63	66	62	79	72		73
23	69	72	65	68	64	80	74		75
24	71	74	66	70	66		77		77
25	73	76	68	71	67		80		79
26	75	78	69	73	69				80
27	77	79	71	75	71				
28	79	80	72	77	73				
29	80		74	79	75				
30			75	80	76				
31			77		78				
32			78		80				
33			80						

Note. AN = Analogies; FW = Figure Weights; IN = Information; MR = Matrix Reasoning; OR = Object Rotation; SO = Spatial Orientation; STV = SToVeS; VP = Visual Puzzles; VO = Vocabulary. Blank cells indicate raw scores beyond RIOT's 31T and 80T hard cutoffs. The actual subtest may contain fewer questions than are needed to reach the 80T ceiling.

Table A.3: Raw-to-T-score conversion, VRI, FRI, and SAI subtests, Age 60+

Raw	VO	IN	AN	MR	VP	FW	OR	STV	SO
0		31						34	31
1		32	31					36	32
2		33	32					38	34
3		35	34	31		31		40	36
4	31	37	35	33		34		43	38
5	33	38	37	35	31	37	31	45	40
6	35	40	38	37	32	39	32	47	42
7	37	42	39	39	35	42	34	49	44
8	38	43	41	41	37	45	37	51	46
9	40	45	42	43	39	48	40	54	48
10	42	47	44	45	41	50	43	56	50
11	44	48	45	47	43	53	45	58	51
12	45	50	46	48	45	56	48	60	53
13	47	52	48	50	47	58	51	63	55
14	49	53	49	52	49	61	53	65	57
15	50	55	51	54	51	64	56	67	59
16	52	57	52	56	53	67	59	69	61
17	54	58	53	58	55	69	62	72	63
18	56	60	55	60	57	72	64	74	65
19	57	62	56	62	59	75	67	76	67
20	59	63	58	64	61	78	70	78	69
21	61	65	59	66	64	80	73	80	71
22	63	67	60	68	66		75		73
23	64	68	62	70	68		78		75
24	66	70	63	72	70		80		76
25	68	72	65	74	72				78
26	70	73	66	76	74				80
27	71	75	67	77	76				
28	73	77	69	79	78				
29	75	78	70	80	80				
30	76	80	71						
31	78		73						
32	80		74						
33			76						
34			77						
35			78						
36			80						

Note. AN = Analogies; FW = Figure Weights; IN = Information; MR = Matrix Reasoning; OR = Object Rotation; SO = Spatial Orientation; STV = SToVeS; VP = Visual Puzzles; VO = Vocabulary. Blank cells indicate raw scores beyond RIOT's 31T and 80T hard cutoffs. The actual subtest may contain fewer questions than are needed to reach the 80T ceiling.

Table A.4: Raw-to-T-score conversion, WMI subtests, by age group

	CS			EM			VR		
Raw	18–39	40–59	60+	18–39	40–59	60+	18–39	40–59	60+
0				33	32	31			
1		31		36	35	36	31		
2	31	32	31	38	38	40	32	31	31
3	35	35	34	41	42	45	35	35	33
4	38	39	38	44	45	49	39	38	37
5	42	42	42	47	48	54	42	42	42
6	45	45	46	50	51	58	46	46	46
7	49	48	49	52	55	63	49	49	51
8	52	52	53	55	58	67	53	53	55
9	56	55	57	58	61	71	56	57	60
10	60	58	61	61	64	76	60	60	64
11	63	62	65	64	68	80	63	64	69
12	67	65	68	66	71		67	68	73
13	70	68	72	69	74		70	71	78
14	74	72	76	72	77		74	75	80
15	78	75	80	75	80		77	79	
16	80	78		78			80	80	
17		80		80					

Note. CS = Computation Span; EM = Exposure Memory; VR = Visual Reversal. Blank cells indicate raw scores beyond RIOT's 31T and 80T hard cutoffs. The actual subtest may contain fewer questions than are needed to reach the 80T ceiling.

Table A.5: Raw-to-T-score conversion, PSI subtests, by age group

	SS			AM				SS			AM		
Time (s)	18–39	40–59	60+	18–39	40–59	60+	Time (s)	18–39	40–59	60+	18–39	40–59	60+
0.10	78						3.00	53	55	58		32	37
0.20	77						3.10	53	54	57		31	35
0.30	76						3.20	52	53	56			33
0.40	75						3.30	51	52	55			32
0.50	74				80	80	3.40	50	51	54			31
0.60	74	80		80	78	78	3.50	49	50	53			
0.70	73	79		78	76	77	3.60	48	49	52			
0.80	72	78		76	74	75	3.70	48	48	51			
0.90	71	77	80	73	72	73	3.80	47	47	50			
1.00	70	76	79	71	70	71	3.90	46	45	49			
1.10	69	75	78	69	68	70	4.00	45	44	48			
1.20	69	74	77	66	66	68	4.10	44	43	47			
1.30	68	73	76	64	65	66	4.20	43	42	46			
1.40	67	72	75	62	63	64	4.30	43	41	45			
1.50	66	71	74	59	61	63	4.40	42	40	44			
1.60	65	70	73	57	59	61	4.50	41	39	43			
1.70	64	69	72	55	57	59	4.60	40	38	42			
1.80	64	68	71	52	55	58	4.70	39	37	41			
1.90	63	67	70	50	53	56	4.80	38	36	40			
2.00	62	65	69	48	51	54	4.90	37	35	39			
2.10	61	64	68	45	49	52	5.00	37	34	38			
2.20	60	63	67	43	48	51	5.10	36	33	37			
2.30	59	62	66	41	46	49	5.20	35	32	36			
2.40	58	61	65	38	44	47	5.30	34	31	35			
2.50	58	60	64	36	42	46	5.40	33		34			
2.60	57	59	63	34	40	44	5.50	32		33			
2.70	56	58	62	31	38	42	5.60	32		32			
2.80	55	57	61		36	40	5.70	31		31			
2.90	54	56	59		34	39							

Note. SS = Symbol Search; AM = Abstract Matching. Blank cells indicate raw scores beyond RIOT's 31T and 80T hard cutoffs. See Section 3.1 for how each subtest's raw score is calculated.

Table A.6: Raw-to-T-score conversion, RTI subtests, by age group

	Simple			Choice				Simple			Choice		
Time (ms)	18–39	40–59	60+	18–39	40–59	60+	Time (ms)	18–39	40–59	60+	18–39	40–59	60+
150	72	71	73	77	80	78	530	34	39	42	50	53	54
160	71	70	73	76	79	77	540	33	38	41	49	52	53
170	70	69	72	76	79	76	550	32	37	40	49	51	53
180	69	68	71	75	78	76	560	31	37	39	48	51	52
190	68	68	70	74	77	75	570		36	38	47	50	51
200	67	67	69	74	76	74	580		35	37	46	49	51
210	66	66	68	73	76	74	590		34	37	46	49	50
220	65	65	68	72	75	73	600		33	36	45	48	49
230	64	64	67	71	74	73	610		32	35	44	47	49
240	63	63	66	71	74	72	620		32	34	44	46	48
250	62	63	65	70	73	71	630		31	33	43	46	47
260	61	62	64	69	72	71	640			32	42	45	47
270	60	61	63	69	71	70	650			32	41	44	46
280	59	60	63	68	71	69	660			31	41	44	46
290	58	59	62	67	70	69	670				40	43	45
300	57	58	61	66	69	68	680				39	42	44
310	56	58	60	66	69	68	690				39	41	44
320	55	57	59	65	68	67	700				38	41	43
330	54	56	58	64	67	66	710				37	40	42
340	53	55	58	64	66	66	720				36	39	42
350	52	54	57	63	66	65	730				36	39	41
360	51	53	56	62	65	64	740				35	38	41
370	50	53	55	61	64	64	750				34	37	40
380	49	52	54	61	64	63	760				34	36	39
390	48	51	53	60	63	63	770				33	36	39
400	47	50	53	59	62	62	780				32	35	38
410	46	49	52	59	61	61	790				31	34	37
420	45	48	51	58	61	61	800				31	34	37
430	44	47	50	57	60	60	810					33	36
440	43	47	49	56	59	59	820					32	36
450	42	46	48	56	59	59	830					31	35
460	41	45	47	55	58	58	840					31	34
470	40	44	47	54	57	58	850						34
480	39	43	46	54	56	57	860						33
490	38	42	45	53	56	56	870						32
500	37	42	44	52	55	56	880						32
510	36	41	43	51	54	55	890						31
520	35	40	42	51	54	54							

Note. Simple and Choice refer to the two timed portions of the Reaction Time subtest (Table 2). Values reflect the test's hard cutoffs for T-scores (31T–80T) and reaction times (150–10,000 ms).

Appendix B T-Score to Composite Lookup Tables

Table B.1 converts the sum of T-scores (ST) across the Basic test's five subtests – Vocabulary, Matrix Reasoning, SToVeS, Visual Reversal, and Symbol Search – directly into an FSIQ. Because RIOT weights subtests by their factor loadings rather than summing them equally, the actual composite for a given ST can vary, as shown in the Possible range column. RIOT also clips the FSIQ to a floor of 71 and a ceiling of 145, so examinees who received a 145 (or 71) can use this table to estimate their unclipped FSIQ.

Table B.1: ST to FSIQ lookup table – Basic Test

ST	FSIQ	Possible range	ST	FSIQ	Possible range	ST	FSIQ	Possible range
155	59	59–59	237	94	89–99	319	130	125–134
156	60	60–60	238	95	90–99	320	130	126–135
157	60	60–60	239	95	90–100	321	130	126–135
158	61	60–61	240	96	91–100	322	131	127–136
159	61	61–61	241	96	91–101	323	131	127–136
160	61	61–62	242	97	91–101	324	132	128–136
161	62	61–62	243	97	92–102	325	132	128–137
162	62	62–63	244	97	92–102	326	133	129–137
163	63	62–63	245	98	92–103	327	133	129–137
164	63	63–64	246	98	93–103	328	133	130–138
165	64	63–64	247	99	93–104	329	134	130–138
166	64	63–65	248	99	94–104	330	134	131–139
167	64	64–65	249	100	94–104	331	135	131–139
168	65	64–66	250	100	94–105	332	135	131–139
169	65	64–66	251	100	95–105	333	136	132–140
170	66	65–67	252	101	95–106	334	136	132–140
171	66	65–67	253	101	95–106	335	136	133–140
172	67	65–68	254	102	96–107	336	137	133–141
173	67	66–68	255	102	96–107	337	137	134–141
174	67	66–68	256	103	97–108	338	138	134–142
175	68	67–69	257	103	97–108	339	138	135–142
176	68	67–69	258	103	98–109	340	139	135–142
177	69	67–70	259	104	98–109	341	139	136–143
178	69	68–70	260	104	99–109	342	139	136–143
179	70	68–71	261	105	99–110	343	140	137–143
180	70	68–71	262	105	99–110	344	140	137–144
181	70	69–72	263	106	100–111	345	141	138–144
182	71	69–72	264	106	100–111	346	141	138–145
183	71	70–73	265	106	101–112	347	142	139–145
184	72	70–73	266	107	101–112	348	142	139–145
185	72	70–74	267	107	102–113	349	142	140–146
186	73	71–74	268	108	102–113	350	143	140–146
187	73	71–75	269	108	103–113	351	143	141–146
188	73	71–75	270	109	103–114	352	144	141–147
189	74	72–76	271	109	103–114	353	144	141–147
190	74	72–76	272	109	104–115	354	145	142–147
191	75	72–77	273	110	104–115	355	145	142–148
192	75	73–77	274	110	105–116	356	145	143–148

continued on next page

Table B.1 – *continued from previous page*

ST	FSIQ	Possible range	ST	FSIQ	Possible range	ST	FSIQ	Possible range
193	76	73–78	275	111	105–116	357	146	143–149
194	76	74–78	276	111	106–117	358	146	144–149
195	76	74–79	277	112	106–117	359	147	144–149
196	77	74–79	278	112	107–117	360	147	145–150
197	77	75–80	279	112	107–118	361	148	145–150
198	78	75–80	280	113	107–118	362	148	146–150
199	78	75–81	281	113	108–119	363	148	146–151
200	79	76–81	282	114	108–119	364	149	147–151
201	79	76–82	283	114	109–120	365	149	147–152
202	79	76–82	284	115	109–120	366	150	148–152
203	80	77–83	285	115	110–121	367	150	148–152
204	80	77–83	286	115	110–121	368	151	149–153
205	81	78–84	287	116	111–121	369	151	149–153
206	81	78–84	288	116	111–122	370	151	150–153
207	82	78–85	289	117	111–122	371	152	150–154
208	82	79–85	290	117	112–123	372	152	151–154
209	82	79–85	291	118	112–123	373	153	151–154
210	83	79–86	292	118	113–124	374	153	152–155
211	83	80–86	293	118	113–124	375	154	152–155
212	84	80–87	294	119	114–125	376	154	153–156
213	84	81–87	295	119	114–125	377	154	153–156
214	85	81–88	296	120	115–125	378	155	154–156
215	85	81–88	297	120	115–126	379	155	154–157
216	85	82–89	298	121	115–126	380	156	155–157
217	86	82–89	299	121	116–127	381	156	155–157
218	86	82–90	300	121	116–127	382	157	156–158
219	87	83–90	301	122	117–128	383	157	156–158
220	87	83–91	302	122	117–128	384	157	157–158
221	88	84–91	303	123	118–129	385	158	157–159
222	88	84–92	304	123	118–129	386	158	158–159
223	88	84–92	305	124	119–129	387	159	158–160
224	89	85–93	306	124	119–130	388	159	159–160
225	89	85–93	307	124	120–130	389	160	159–160
226	90	85–94	308	125	120–130	390	160	159–161
227	90	86–94	309	125	121–131	391	160	160–161
228	91	86–94	310	126	121–131	392	161	160–161
229	91	87–95	311	126	121–131	393	161	161–162
230	91	87–95	312	127	122–132	394	162	161–162
231	92	87–96	313	127	122–132	395	162	162–163
232	92	88–96	314	127	123–133	396	163	162–163
233	93	88–97	315	128	123–133	397	163	163–163
234	93	88–97	316	128	124–133	398	163	163–164
235	94	89–98	317	129	124–134	399	164	164–164
236	94	89–98	318	129	125–134	400	164	164–164

Note. Not every subtest may reach 80T in practice, so the true maximum is likely somewhat lower.

Table B.2 converts the ST across the Full test's sixteen subscores into an FSIQ. Because RIOT weights subtests by their factor loadings rather than summing them equally, the actual composite for a given ST can vary, as shown in the Possible range column. RIOT also clips the FSIQ to a floor of 71 and a ceiling of 145, so examinees who received a 145 (or 71) can use this table to estimate their unclipped FSIQ.

Table B.2: ST to FSIQ lookup table – Full Test

ST	FSIQ	Possible range	ST	FSIQ	Possible range	ST	FSIQ	Possible range
496	56	56–56	758	94	85–102	1020	132	124–141
497	56	56–56	759	94	85–102	1021	132	124–141
498	56	56–56	760	94	85–102	1022	132	124–141
499	56	56–56	761	94	85–103	1023	133	124–141
500	56	56–56	762	94	85–103	1024	133	125–142
501	56	56–57	763	95	86–103	1025	133	125–142
502	57	56–57	764	95	86–103	1026	133	125–142
503	57	56–57	765	95	86–103	1027	133	125–142
504	57	56–57	766	95	86–103	1028	133	125–142
505	57	56–57	767	95	86–104	1029	133	125–142
506	57	57–58	768	95	86–104	1030	134	126–142
507	57	57–58	769	95	86–104	1031	134	126–142
508	57	57–58	770	96	86–104	1032	134	126–143
509	58	57–58	771	96	87–104	1033	134	126–143
510	58	57–59	772	96	87–104	1034	134	126–143
511	58	57–59	773	96	87–104	1035	134	126–143
512	58	57–59	774	96	87–105	1036	134	127–143
513	58	57–59	775	96	87–105	1037	135	127–143
514	58	57–59	776	96	87–105	1038	135	127–143
515	58	57–60	777	97	87–105	1039	135	127–143
516	59	57–60	778	97	88–105	1040	135	127–144
517	59	58–60	779	97	88–105	1041	135	127–144
518	59	58–60	780	97	88–106	1042	135	128–144
519	59	58–60	781	97	88–106	1043	135	128–144
520	59	58–61	782	97	88–106	1044	136	128–144
521	59	58–61	783	98	88–106	1045	136	128–144
522	59	58–61	784	98	88–106	1046	136	128–144
523	60	58–61	785	98	88–106	1047	136	128–144
524	60	58–61	786	98	89–107	1048	136	129–145
525	60	58–62	787	98	89–107	1049	136	129–145
526	60	58–62	788	98	89–107	1050	136	129–145
527	60	58–62	789	98	89–107	1051	137	129–145
528	60	59–62	790	99	89–107	1052	137	129–145
529	60	59–62	791	99	89–107	1053	137	129–145
530	61	59–63	792	99	89–108	1054	137	130–145
531	61	59–63	793	99	89–108	1055	137	130–145
532	61	59–63	794	99	90–108	1056	137	130–146
533	61	59–63	795	99	90–108	1057	138	130–146
534	61	59–63	796	99	90–108	1058	138	130–146
535	61	59–64	797	100	90–108	1059	138	130–146
536	61	59–64	798	100	90–109	1060	138	131–146

continued on next page

Table B.2 – *continued from previous page*

ST	FSIQ	Possible range	ST	FSIQ	Possible range	ST	FSIQ	Possible range
537	62	59–64	799	100	90–109	1061	138	131–146
538	62	59–64	800	100	90–109	1062	138	131–146
539	62	59–64	801	100	91–109	1063	138	131–146
540	62	60–65	802	100	91–109	1064	139	131–147
541	62	60–65	803	100	91–109	1065	139	131–147
542	62	60–65	804	101	91–109	1066	139	132–147
543	62	60–65	805	101	91–110	1067	139	132–147
544	63	60–66	806	101	91–110	1068	139	132–147
545	63	60–66	807	101	91–110	1069	139	132–147
546	63	60–66	808	101	91–110	1070	139	132–147
547	63	60–66	809	101	92–110	1071	140	132–147
548	63	60–66	810	101	92–110	1072	140	133–148
549	63	60–66	811	102	92–111	1073	140	133–148
550	64	61–67	812	102	92–111	1074	140	133–148
551	64	61–67	813	102	92–111	1075	140	133–148
552	64	61–67	814	102	92–111	1076	140	133–148
553	64	61–67	815	102	92–111	1077	140	133–148
554	64	61–67	816	102	93–111	1078	141	133–148
555	64	61–68	817	102	93–112	1079	141	134–148
556	64	61–68	818	103	93–112	1080	141	134–149
557	65	61–68	819	103	93–112	1081	141	134–149
558	65	61–68	820	103	93–112	1082	141	134–149
559	65	61–68	821	103	93–112	1083	141	134–149
560	65	62–68	822	103	93–112	1084	141	134–149
561	65	62–69	823	103	93–112	1085	142	135–149
562	65	62–69	824	104	94–113	1086	142	135–149
563	65	62–69	825	104	94–113	1087	142	135–149
564	66	62–69	826	104	94–113	1088	142	135–150
565	66	62–69	827	104	94–113	1089	142	135–150
566	66	62–69	828	104	94–113	1090	142	135–150
567	66	62–70	829	104	94–113	1091	142	136–150
568	66	62–70	830	104	94–114	1092	143	136–150
569	66	62–70	831	105	94–114	1093	143	136–150
570	66	63–70	832	105	95–114	1094	143	136–150
571	67	63–70	833	105	95–114	1095	143	136–150
572	67	63–71	834	105	95–114	1096	143	136–150
573	67	63–71	835	105	95–114	1097	143	137–151
574	67	63–71	836	105	95–115	1098	143	137–151
575	67	63–71	837	105	95–115	1099	144	137–151
576	67	63–71	838	106	95–115	1100	144	137–151
577	67	63–71	839	106	96–115	1101	144	137–151
578	68	63–72	840	106	96–115	1102	144	138–151
579	68	63–72	841	106	96–115	1103	144	138–151
580	68	64–72	842	106	96–116	1104	144	138–151
581	68	64–72	843	106	96–116	1105	145	138–152
582	68	64–72	844	106	96–116	1106	145	138–152
583	68	64–73	845	107	96–116	1107	145	138–152
584	68	64–73	846	107	97–116	1108	145	139–152

continued on next page

Table B.2 – *continued from previous page*

ST	FSIQ	Possible range	ST	FSIQ	Possible range	ST	FSIQ	Possible range
585	69	64–73	847	107	97–116	1109	145	139–152
586	69	64–73	848	107	97–116	1110	145	139–152
587	69	64–73	849	107	97–117	1111	145	139–152
588	69	64–73	850	107	97–117	1112	146	139–152
589	69	65–74	851	107	97–117	1113	146	139–153
590	69	65–74	852	108	98–117	1114	146	140–153
591	69	65–74	853	108	98–117	1115	146	140–153
592	70	65–74	854	108	98–117	1116	146	140–153
593	70	65–74	855	108	98–118	1117	146	140–153
594	70	65–75	856	108	98–118	1118	146	140–153
595	70	65–75	857	108	98–118	1119	147	140–153
596	70	65–75	858	108	98–118	1120	147	141–153
597	70	65–75	859	109	99–118	1121	147	141–154
598	71	65–75	860	109	99–118	1122	147	141–154
599	71	66–75	861	109	99–118	1123	147	141–154
600	71	66–76	862	109	99–119	1124	147	141–154
601	71	66–76	863	109	99–119	1125	147	141–154
602	71	66–76	864	109	99–119	1126	148	142–154
603	71	66–76	865	109	99–119	1127	148	142–154
604	71	66–76	866	110	100–119	1128	148	142–154
605	72	66–76	867	110	100–119	1129	148	142–155
606	72	66–77	868	110	100–120	1130	148	142–155
607	72	67–77	869	110	100–120	1131	148	142–155
608	72	67–77	870	110	100–120	1132	148	143–155
609	72	67–77	871	110	100–120	1133	149	143–155
610	72	67–77	872	111	101–120	1134	149	143–155
611	72	67–77	873	111	101–120	1135	149	143–155
612	73	67–78	874	111	101–121	1136	149	143–155
613	73	67–78	875	111	101–121	1137	149	143–155
614	73	67–78	876	111	101–121	1138	149	144–156
615	73	67–78	877	111	101–121	1139	149	144–156
616	73	68–78	878	111	101–121	1140	150	144–156
617	73	68–78	879	112	102–121	1141	150	144–156
618	73	68–79	880	112	102–121	1142	150	144–156
619	74	68–79	881	112	102–122	1143	150	144–156
620	74	68–79	882	112	102–122	1144	150	145–156
621	74	68–79	883	112	102–122	1145	150	145–156
622	74	68–79	884	112	102–122	1146	150	145–157
623	74	68–80	885	112	102–122	1147	151	145–157
624	74	68–80	886	113	103–122	1148	151	145–157
625	74	69–80	887	113	103–123	1149	151	145–157
626	75	69–80	888	113	103–123	1150	151	146–157
627	75	69–80	889	113	103–123	1151	151	146–157
628	75	69–80	890	113	103–123	1152	151	146–157
629	75	69–81	891	113	103–123	1153	152	146–157
630	75	69–81	892	113	104–123	1154	152	146–157
631	75	69–81	893	114	104–123	1155	152	147–158
632	75	69–81	894	114	104–124	1156	152	147–158

continued on next page

Table B.2 – *continued from previous page*

ST	FSIQ	Possible range	ST	FSIQ	Possible range	ST	FSIQ	Possible range
633	76	70–81	895	114	104–124	1157	152	147–158
634	76	70–81	896	114	104–124	1158	152	147–158
635	76	70–82	897	114	104–124	1159	152	147–158
636	76	70–82	898	114	105–124	1160	153	147–158
637	76	70–82	899	114	105–124	1161	153	148–158
638	76	70–82	900	115	105–125	1162	153	148–158
639	77	70–82	901	115	105–125	1163	153	148–158
640	77	70–82	902	115	105–125	1164	153	148–159
641	77	70–83	903	115	105–125	1165	153	148–159
642	77	71–83	904	115	105–125	1166	153	148–159
643	77	71–83	905	115	106–125	1167	154	149–159
644	77	71–83	906	115	106–125	1168	154	149–159
645	77	71–83	907	116	106–126	1169	154	149–159
646	78	71–83	908	116	106–126	1170	154	149–159
647	78	71–84	909	116	106–126	1171	154	149–159
648	78	71–84	910	116	106–126	1172	154	149–160
649	78	71–84	911	116	107–126	1173	154	150–160
650	78	72–84	912	116	107–126	1174	155	150–160
651	78	72–84	913	116	107–127	1175	155	150–160
652	78	72–84	914	117	107–127	1176	155	150–160
653	79	72–85	915	117	107–127	1177	155	150–160
654	79	72–85	916	117	107–127	1178	155	150–160
655	79	72–85	917	117	108–127	1179	155	151–160
656	79	72–85	918	117	108–127	1180	155	151–160
657	79	72–85	919	117	108–127	1181	156	151–161
658	79	72–85	920	118	108–128	1182	156	151–161
659	79	73–86	921	118	108–128	1183	156	151–161
660	80	73–86	922	118	108–128	1184	156	152–161
661	80	73–86	923	118	108–128	1185	156	152–161
662	80	73–86	924	118	109–128	1186	156	152–161
663	80	73–86	925	118	109–128	1187	156	152–161
664	80	73–87	926	118	109–128	1188	157	152–161
665	80	73–87	927	119	109–129	1189	157	152–161
666	80	73–87	928	119	109–129	1190	157	153–161
667	81	74–87	929	119	109–129	1191	157	153–162
668	81	74–87	930	119	110–129	1192	157	153–162
669	81	74–87	931	119	110–129	1193	157	153–162
670	81	74–88	932	119	110–129	1194	158	153–162
671	81	74–88	933	119	110–130	1195	158	154–162
672	81	74–88	934	120	110–130	1196	158	154–162
673	81	74–88	935	120	110–130	1197	158	154–162
674	82	74–88	936	120	110–130	1198	158	154–162
675	82	75–88	937	120	111–130	1199	158	154–162
676	82	75–89	938	120	111–130	1200	158	154–162
677	82	75–89	939	120	111–130	1201	159	155–163
678	82	75–89	940	120	111–131	1202	159	155–163
679	82	75–89	941	121	111–131	1203	159	155–163
680	82	75–89	942	121	111–131	1204	159	155–163

continued on next page

Table B.2 – *continued from previous page*

ST	FSIQ	Possible range	ST	FSIQ	Possible range	ST	FSIQ	Possible range
681	83	75–89	943	121	112–131	1205	159	155–163
682	83	75–90	944	121	112–131	1206	159	155–163
683	83	76–90	945	121	112–131	1207	159	156–163
684	83	76–90	946	121	112–131	1208	160	156–163
685	83	76–90	947	121	112–131	1209	160	156–163
686	83	76–90	948	122	112–132	1210	160	156–164
687	84	76–90	949	122	113–132	1211	160	156–164
688	84	76–91	950	122	113–132	1212	160	157–164
689	84	76–91	951	122	113–132	1213	160	157–164
690	84	76–91	952	122	113–132	1214	160	157–164
691	84	77–91	953	122	113–132	1215	161	157–164
692	84	77–91	954	122	113–132	1216	161	157–164
693	84	77–91	955	123	114–133	1217	161	157–164
694	85	77–92	956	123	114–133	1218	161	158–164
695	85	77–92	957	123	114–133	1219	161	158–164
696	85	77–92	958	123	114–133	1220	161	158–165
697	85	77–92	959	123	114–133	1221	161	158–165
698	85	77–92	960	123	114–133	1222	162	158–165
699	85	78–92	961	123	114–133	1223	162	159–165
700	85	78–93	962	124	115–133	1224	162	159–165
701	86	78–93	963	124	115–134	1225	162	159–165
702	86	78–93	964	124	115–134	1226	162	159–165
703	86	78–93	965	124	115–134	1227	162	159–165
704	86	78–93	966	124	115–134	1228	162	159–165
705	86	78–93	967	124	115–134	1229	163	160–165
706	86	78–94	968	125	116–134	1230	163	160–166
707	86	79–94	969	125	116–134	1231	163	160–166
708	87	79–94	970	125	116–134	1232	163	160–166
709	87	79–94	971	125	116–135	1233	163	160–166
710	87	79–94	972	125	116–135	1234	163	161–166
711	87	79–94	973	125	116–135	1235	163	161–166
712	87	79–94	974	125	117–135	1236	164	161–166
713	87	79–95	975	126	117–135	1237	164	161–166
714	87	79–95	976	126	117–135	1238	164	161–166
715	88	80–95	977	126	117–135	1239	164	162–166
716	88	80–95	978	126	117–136	1240	164	162–166
717	88	80–95	979	126	117–136	1241	164	162–167
718	88	80–95	980	126	117–136	1242	165	162–167
719	88	80–96	981	126	118–136	1243	165	162–167
720	88	80–96	982	127	118–136	1244	165	163–167
721	88	80–96	983	127	118–136	1245	165	163–167
722	89	80–96	984	127	118–136	1246	165	163–167
723	89	81–96	985	127	118–136	1247	165	163–167
724	89	81–96	986	127	118–137	1248	165	163–167
725	89	81–97	987	127	119–137	1249	166	164–167
726	89	81–97	988	127	119–137	1250	166	164–167
727	89	81–97	989	128	119–137	1251	166	164–167
728	89	81–97	990	128	119–137	1252	166	164–168

continued on next page

Table B.2 – *continued from previous page*

ST	FSIQ	Possible range	ST	FSIQ	Possible range	ST	FSIQ	Possible range
729	90	81–97	991	128	119–137	1253	166	165–168
730	90	81–97	992	128	119–137	1254	166	165–168
731	90	81–98	993	128	120–138	1255	166	165–168
732	90	82–98	994	128	120–138	1256	167	165–168
733	90	82–98	995	128	120–138	1257	167	165–168
734	90	82–98	996	129	120–138	1258	167	166–168
735	91	82–98	997	129	120–138	1259	167	166–168
736	91	82–98	998	129	120–138	1260	167	166–168
737	91	82–99	999	129	121–138	1261	167	166–168
738	91	82–99	1000	129	121–138	1262	167	166–168
739	91	82–99	1001	129	121–139	1263	168	167–169
740	91	83–99	1002	129	121–139	1264	168	167–169
741	91	83–99	1003	130	121–139	1265	168	167–169
742	92	83–99	1004	130	121–139	1266	168	167–169
743	92	83–100	1005	130	122–139	1267	168	167–169
744	92	83–100	1006	130	122–139	1268	168	168–169
745	92	83–100	1007	130	122–139	1269	168	168–169
746	92	83–100	1008	130	122–139	1270	169	168–169
747	92	84–100	1009	131	122–140	1271	169	168–169
748	92	84–100	1010	131	122–140	1272	169	168–169
749	93	84–101	1011	131	123–140	1273	169	169–169
750	93	84–101	1012	131	123–140	1274	169	169–170
751	93	84–101	1013	131	123–140	1275	169	169–170
752	93	84–101	1014	131	123–140	1276	169	169–170
753	93	84–101	1015	131	123–140	1277	170	169–170
754	93	84–101	1016	132	123–140	1278	170	170–170
755	93	85–102	1017	132	123–141	1279	170	170–170
756	94	85–102	1018	132	124–141	1280	170	170–170
757	94	85–102	1019	132	124–141			

Note. Not every subtest may reach 80T in practice, so the true maximum is likely somewhat lower.

Table B.3 converts ST into each of the six index composites (Table 1). RIOT clips each index to a floor of 31T (71 on the IQ metric) and a ceiling of 80T (145), so examinees who received an 80T (or 31T) on an index can use this table to estimate their unclipped score.

Table B.3: ST to index composite lookup table

ST	VRI		FRI		SAI		WMI		PSI		RTI	
	T	IQ	T	IQ	T	IQ	T	IQ	T	IQ	T	IQ
62									28	67	30	69
63									29	68	30	70
64									29	69	31	71
65									30	70	31	72
66									30	71	32	73
67									31	71	32	73
68									32	72	33	74
69									32	73	33	75
70									33	74	34	76
71									33	75	34	77
72									34	76	35	77
73									34	77	35	78
74									35	77	36	79
75									36	78	37	80
76									36	79	37	81
77									37	80	38	81
78									37	81	38	82
79									38	82	39	83
80									38	83	39	84
81									39	84	40	85
82									40	84	40	85
83									40	85	41	86
84									41	86	41	87
85									41	87	42	88
86									42	88	42	89
87									42	89	43	90
88									43	90	44	90
89									44	90	44	91
90									44	91	45	92
91									45	92	45	93
92									45	93	46	94
93	28	66	27	66	28	66	23	60	46	94	46	94
94	28	67	28	67	28	67	24	61	47	95	47	95
95	28	68	28	67	28	68	24	62	47	96	47	96
96	29	68	29	68	29	68	25	62	48	97	48	97
97	29	69	29	69	29	69	25	63	48	97	48	98
98	30	69	29	69	30	69	26	64	49	98	49	98
99	30	70	30	70	30	70	26	64	49	99	49	99
100	30	71	30	70	30	70	27	65	50	100	50	100
101	31	71	31	71	31	71	27	66	51	101	51	101
102	31	72	31	72	31	72	28	66	51	102	51	102

continued on next page

Table B.3 – *continued from previous page*

ST	VRI		FRI		SAI		WMI		PSI		RTI	
	T	IQ	T	IQ	T	IQ	T	IQ	T	IQ	T	IQ
103	32	72	31	72	31	72	28	67	52	103	52	102
104	32	73	32	73	32	73	29	68	52	103	52	103
105	32	74	32	73	32	73	29	69	53	104	53	104
106	33	74	33	74	33	74	29	69	53	105	53	105
107	33	75	33	75	33	75	30	70	54	106	54	106
108	34	75	33	75	33	75	30	71	55	107	54	106
109	34	76	34	76	34	76	31	71	55	108	55	107
110	34	76	34	76	34	76	31	72	56	109	55	108
111	35	77	35	77	35	77	32	73	56	110	56	109
112	35	78	35	77	35	78	32	73	57	110	56	110
113	35	78	35	78	35	78	33	74	58	111	57	110
114	36	79	36	79	36	79	33	75	58	112	58	111
115	36	79	36	79	36	79	34	76	59	113	58	112
116	37	80	37	80	37	80	34	76	59	114	59	113
117	37	81	37	80	37	81	35	77	60	115	59	114
118	37	81	37	81	37	81	35	78	60	116	60	115
119	38	82	38	82	38	82	36	78	61	116	60	115
120	38	82	38	82	38	82	36	79	62	117	61	116
121	39	83	39	83	39	83	36	80	62	118	61	117
122	39	84	39	83	39	83	37	80	63	119	62	118
123	39	84	39	84	39	84	37	81	63	120	62	119
124	40	85	40	85	40	85	38	82	64	121	63	119
125	40	85	40	85	40	85	38	83	64	122	63	120
126	41	86	41	86	41	86	39	83	65	123	64	121
127	41	86	41	86	41	86	39	84	66	123	65	122
128	41	87	41	87	41	87	40	85	66	124	65	123
129	42	88	42	88	42	88	40	85	67	125	66	123
130	42	88	42	88	42	88	41	86	67	126	66	124
131	43	89	42	89	43	89	41	87	68	127	67	125
132	43	89	43	89	43	89	42	87	68	128	67	126
133	43	90	43	90	43	90	42	88	69	129	68	127
134	44	91	44	91	44	91	43	89	70	129	68	127
135	44	91	44	91	44	91	43	90	70	130	69	128
136	45	92	44	92	44	92	43	90	71	131	69	129
137	45	92	45	92	45	92	44	91	71	132	70	130
138	45	93	45	93	45	93	44	92	72	133	70	131
139	46	94	46	93	46	94	45	92	73	134	71	131
140	46	94	46	94	46	94	45	93	73	135	71	132
141	46	95	46	95	46	95	46	94	74	135	72	133
142	47	95	47	95	47	95	46	94	74	136	73	134
143	47	96	47	96	47	96	47	95	75	137	73	135
144	48	96	48	96	48	96	47	96	75	138	74	135
145	48	97	48	97	48	97	48	97	76	139	74	136
146	48	98	48	98	48	98	48	97	77	140	75	137
147	49	98	49	98	49	98	49	98	77	141	75	138
148	49	99	49	99	49	99	49	99	78	142	76	139

continued on next page

Table B.3 – *continued from previous page*

ST	VRI		FRI		SAI		WMI		PSI		RTI	
	T	IQ	T	IQ	T	IQ	T	IQ	T	IQ	T	IQ
149	50	99	50	99	50	99	50	99	78	142	76	140
150	50	100	50	100	50	100	50	100	79	143	77	140
151	50	101	50	101	50	101	50	101	79	144	77	141
152	51	101	51	101	51	101	51	101	80	145	78	142
153	51	102	51	102	51	102	51	102	81	146	78	143
154	52	102	52	102	52	102	52	103	81	147	79	144
155	52	103	52	103	52	103	52	103	82	148	80	144
156	52	104	52	104	52	104	53	104	82	148	80	145
157	53	104	53	104	53	104	53	105	83	149	81	146
158	53	105	53	105	53	105	54	106	83	150	81	147
159	54	105	54	105	54	105	54	106	84	151	82	148
160	54	106	54	106	54	106	55	107	85	152	82	148
161	54	106	54	107	54	106	55	108				
162	55	107	55	107	55	107	56	108				
163	55	108	55	108	55	108	56	109				
164	55	108	56	108	56	108	57	110				
165	56	109	56	109	56	109	57	110				
166	56	109	56	109	56	109	57	111				
167	57	110	57	110	57	110	58	112				
168	57	111	57	111	57	111	58	113				
169	57	111	58	111	57	111	59	113				
170	58	112	58	112	58	112	59	114				
171	58	112	58	112	58	112	60	115				
172	59	113	59	113	59	113	60	115				
173	59	114	59	114	59	114	61	116				
174	59	114	59	114	59	114	61	117				
175	60	115	60	115	60	115	62	117				
176	60	115	60	115	60	115	62	118				
177	61	116	61	116	61	116	63	119				
178	61	116	61	117	61	117	63	120				
179	61	117	61	117	61	117	64	120				
180	62	118	62	118	62	118	64	121				
181	62	118	62	118	62	118	64	122				
182	63	119	63	119	63	119	65	122				
183	63	119	63	120	63	119	65	123				
184	63	120	63	120	63	120	66	124				
185	64	121	64	121	64	121	66	124				
186	64	121	64	121	64	121	67	125				
187	65	122	65	122	65	122	67	126				
188	65	122	65	123	65	122	68	127				
189	65	123	65	123	65	123	68	127				
190	66	124	66	124	66	124	69	128				
191	66	124	66	124	66	124	69	129				
192	66	125	67	125	67	125	70	129				
193	67	125	67	125	67	125	70	130				
194	67	126	67	126	67	126	71	131				

continued on next page

Table B.3 – *continued from previous page*

ST	VRI		FRI		SAI		WMI		PSI		RTI	
	T	IQ	T	IQ	T	IQ	T	IQ	T	IQ	T	IQ
195	68	126	68	127	68	127	71	131				
196	68	127	68	127	68	127	71	132				
197	68	128	69	128	69	128	72	133				
198	69	128	69	128	69	128	72	134				
199	69	129	69	129	69	129	73	134				
200	70	129	70	130	70	130	73	135				
201	70	130	70	130	70	130	74	136				
202	70	131	71	131	70	131	74	136				
203	71	131	71	131	71	131	75	137				
204	71	132	71	132	71	132	75	138				
205	72	132	72	133	72	132	76	138				
206	72	133	72	133	72	133	76	139				
207	72	134	73	134	72	134	77	140				
208	73	134	73	134	73	134	77	141				
209	73	135	73	135	73	135	78	141				
210	74	135	74	136	74	135	78	142				
211	74	136	74	136	74	136	78	143				
212	74	136	75	137	74	137	79	143				
213	75	137	75	137	75	137	79	144				
214	75	138	75	138	75	138	80	145				
215	75	138	76	139	76	138	80	145				
216	76	139	76	139	76	139	81	146				
217	76	139	76	140	76	140	81	147				
218	77	140	77	140	77	140	82	148				
219	77	141	77	141	77	141	82	148				
220	77	141	78	142	78	141	83	149				
221	78	142	78	142	78	142	83	150				
222	78	142	78	143	78	143	84	150				
223	79	143	79	143	79	143	84	151				
224	79	144	79	144	79	144	85	152				
225	79	144	80	144	80	144	85	152				
226	80	145	80	145	80	145	85	153				
227	80	145	80	146	80	145	86	154				
228	81	146	81	146	81	146	86	155				
229	81	146	81	147	81	147	87	155				
230	81	147	82	147	82	147	87	156				
231	82	148	82	148	82	148	88	157				
232	82	148	82	149	82	148	88	157				
233	83	149	83	149	83	149	89	158				
234	83	149	83	150	83	150	89	159				
235	83	150	84	150	83	150	90	159				
236	84	151	84	151	84	151	90	160				
237	84	151	84	152	84	151	91	161				
238	85	152	85	152	85	152	91	162				
239	85	152	85	153	85	153	92	162				
240	85	153	86	153	85	153	92	163				

Note. Not every subtest may reach 80T in practice, so the true maximum is likely somewhat lower.

Appendix C Full `lavaan` Model Output

Figure C.1: Full `lavaan` output, reproduced hierarchical six-factor model

```
lavaan 0.6-20 ended normally after 72 iterations

  Estimator                                         ML
  Optimization method                           NLMINB
  Number of model parameters                        38

  Number of observations                           417

Model Test User Model:

  Test statistic                               238.972
  Degrees of freedom                               98
  P-value (Chi-square)                          0.000

Model Test Baseline Model:

  Test statistic                              2858.815
  Degrees of freedom                              120
  P-value                                       0.000

User Model versus Baseline Model:

  Comparative Fit Index (CFI)                   0.949
  Tucker-Lewis Index (TLI)                      0.937

Loglikelihood and Information Criteria:

  Loglikelihood user model (H0)             -8149.227
  Loglikelihood unrestricted model (H1)     -8029.741

  Akaike (AIC)                              16374.454
  Bayesian (BIC)                            16527.711
  Sample-size adjusted Bayesian (SABIC)     16407.127

Root Mean Square Error of Approximation:

  RMSEA                                         0.059
  90 Percent confidence interval - lower        0.049
  90 Percent confidence interval - upper        0.068
  P-value H_0: RMSEA <= 0.050                   0.063
  P-value H_0: RMSEA >= 0.080                   0.000

Standardized Root Mean Square Residual:

  SRMR                                          0.048

Parameter Estimates:

  Standard errors                            Standard
  Information                                Expected
  Information saturated (h1) model          Structured

Latent Variables:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
```

```
  VRI =~
    VO                0.512    0.037   13.779    0.000    0.789    0.790
    IN                0.499    0.037   13.568    0.000    0.769    0.770
    AN                0.475    0.036   13.076    0.000    0.732    0.733
  FRI =~
    MR                0.305    0.042    7.359    0.000    0.766    0.767
    VP                0.285    0.039    7.268    0.000    0.715    0.716
    FW                0.310    0.042    7.370    0.000    0.777    0.778
  SAI =~
    STV               0.146    0.067    2.181    0.029    0.716    0.717
    OR                0.141    0.065    2.180    0.029    0.695    0.696
    SO                0.172    0.079    2.172    0.030    0.845    0.846
  WMI =~
    CS                0.180    0.066    2.743    0.006    0.495    0.496
    EM                0.160    0.059    2.720    0.007    0.440    0.440
    VR                0.214    0.078    2.750    0.006    0.589    0.590
  PSI =~
    SS                0.687    0.070    9.834    0.000    0.845    0.846
    AM                0.461    0.045   10.240    0.000    0.567    0.568
  RTI =~
    SRT               0.744    0.047   15.747    0.000    0.864    0.865
    CRT               0.725    0.046   15.803    0.000    0.843    0.844
  g =~
    VRI               1.173    0.114   10.317    0.000    0.761    0.761
    FRI               2.299    0.345    6.653    0.000    0.917    0.917
    SAI               4.820    2.279    2.115    0.034    0.979    0.979
    WMI               2.559    0.950    2.694    0.007    0.931    0.931
    PSI               0.716    0.095    7.515    0.000    0.582    0.582
    RTI               0.592    0.069    8.520    0.000    0.509    0.509

Variances:
                   Estimate  Std.Err  z-value  P(>|z|)   Std.lv  Std.all
   .VO                0.375    0.039    9.685    0.000    0.375    0.376
   .IN                0.406    0.040   10.258    0.000    0.406    0.407
   .AN                0.461    0.041   11.134    0.000    0.461    0.463
   .MR                0.411    0.037   11.118    0.000    0.411    0.412
   .VP                0.486    0.040   12.030    0.000    0.486    0.487
   .FW                0.393    0.036   10.841    0.000    0.393    0.394
   .STV               0.484    0.039   12.452    0.000    0.484    0.485
   .OR                0.514    0.041   12.686    0.000    0.514    0.515
   .SO                0.284    0.031    9.289    0.000    0.284    0.285
   .CS                0.752    0.058   12.883    0.000    0.752    0.754
   .EM                0.804    0.060   13.383    0.000    0.804    0.806
   .VR                0.650    0.057   11.437    0.000    0.650    0.652
   .SS                0.284    0.092    3.075    0.002    0.284    0.285
   .AM                0.676    0.062   10.901    0.000    0.676    0.678
   .SRT               0.251    0.062    4.071    0.000    0.251    0.252
   .CRT               0.287    0.060    4.813    0.000    0.287    0.288
   .VRI               1.000                              0.421    0.421
   .FRI               1.000                              0.159    0.159
   .SAI               1.000                              0.041    0.041
   .WMI               1.000                              0.132    0.132
   .PSI               1.000                              0.661    0.661
   .RTI               1.000                              0.741    0.741
    g                 1.000                              1.000    1.000
```

Note. Model discussed in Section 8.1. Includes unstandardized estimates, standard errors, z-values, and all fit indices.