# Confirmatory Factor Analysis and Multiple Regression Analysis With the Shipley-2, WASI, and WIAT-II

Jesse R. Pace and Matthew R. Reynolds
University of Kansas

## Abstract

This paper explored the combined factor structure of the Shipley-2 and the Wechsler Abbreviated Scales of Intelligence (WASI) to help determine what the Shipley-2 measures, and also analyzed the utility of the Shipley-2 in predicting academic test scores on the Wechsler Individual Achievement Test, Second Edition (WIAT-II). Data from the Shipley-2's concurrent validity sample with the WASI ($n$ = 100) was used in cross-battery confirmatory factor analysis (CFA) and multiple regression analysis. Cross-battery CFA was used to load the Shipley-2 and WASI subtests to the same factors in order to determine if a three, two, or one factor model best fit the data. Based on subsequent fit indices of these models, the Shipley-2 is better explained by a three factor (Gc, Gf, Gv) model than a two (Gc, Gf/Gv) or a one factor (Gc/Gf/Gv) model. Multiple regression analysis compared the amount of unique variance explained by the Shipley-2, after accounting for variance explained by the WASI, on the WIAT-II's Word Reading and Numerical Operations subtests. The results of this analysis indicated that both tests explain statistically significant variance in the WIAT-II subtests over the other. Overall, the results of the present study lead us to conclude that the Shipley-2 measures three distinct cognitive abilities (Gv, Gf and Gc) instead of only the two posited in the Shipley-2 manual. Also, it appears that the Shipley-2 and WASI should be used as complimentary measures when predicting academic achievement rather than as interchangeable measures.

## Introduction

- The Shipley-2 is a brief measure of intelligence, designed to provide an efficient estimate of overall cognitive ability based on measures of crystalized ability (Gc) and fluid cognitive ability (Gf).
- The Shipley-2 features a new subtest: Block Patterns, which is theorized to measure fluid reasoning.
  - The construct measured by the Block Patterns subtest resembles the Wechsler Block Design subtest.
  - The Block Design task may be a better measure of visual spatial ability (Gv) than a measure of fluid ability (e.g., Chen, Keith, Chen and Chang 2009).
  - The present research sought to determine if the Block Patterns test fits better with tests of fluid reasoning, or with tests of visual-spatial processing.
- Intelligence test scores have been shown to be highly predictive of academic achievement (e.g., Jensen, 1998).
  - Mayes, Calhoun, Bixler, and Zimmerman (2009) found that the Wechsler Abbreviated Scales of Intelligence (WASI, Wechsler, 1999) measure of overall cognitive functioning (FSIQ) explained 35% of the variance in reading achievement, and 22% of the variance in math scores on the Wide Range Achievement Test.
  - The present study sought to determine if the Shipley-2 provides utility in explaining variance academic test scores beyond what is explained by the WASI.

## Method

**Participants**

- Participants came from the Shipley-2's concurrent validity studies. The sample for our confirmatory factor analysis (CFA) analysis was composed of 100 adults from a forensic setting; these same adults were used for the multiple regression analysis, but there was one less participant (99).
- The ethnic/racial makeup of the sample was 63% White, 30% Black, 3% Hispanic, and 4% other. The mean age of the sample was 31.46.
- The average IQ scores of the sample were considerably lower than the average population (see Table 1).
- Three scores were dropped from the multiple regression analysis due to WASI composite-subscale discrepancies (e.g., FSIQ = 98, VCI = 68 and PRI = 58).

**Measurement Instruments**

- Two brief measures of intellectual functioning were used in our analysis: the Shipley-2, and the WASI. A measure of academic achievement was also used: the Wechsler Individual Achievement Test (WIAT-II).
- The Shipley-2's subtests and its two composites, Composite A (i.e. Abstraction and Vocabulary) and Composite B (i.e. Block Patterns and Vocabulary) were used in analysis. From the norming sample, median interrelatedness of the subtest ranged from .77 (Abstraction) to .91 (Block Patterns). Median internal consistency estimates of the composites were .91 for A and .93 for B (Shipley et al., 2009).
- The WASI Full Scale IQ (FSIQ) and its composing subtests, Vocabulary, Block Design, Similarities, and Matrix Reasoning, were also used in analysis. From the WASI norming sample, average internal consistency estimates of the subtests were .92 (Similarities and Block Design) and .94 (Vocabulary and Matrix Reasoning), and .96 for the FSIQ (Wechsler, 1999, as cited in Psychology Board of Australia, 2014)
- Two of the WIAT-II's subtests, Word Reading and Numerical Operations, were used in analysis. From the norming sample, internal consistency estimates were .91 for Numerical Operations, and .97 for Word Reading (Wechsler, 2001).

**Procedure**

- CFA, and multiple regression analysis were used.

1. CFA model design
   Cross battery-CFA was conducted between the Shipley-2 and the WASI. Three models were compared for overall fit: a three factor Gf, Gv, and Gc model, a two factor Gf/Gv and Gc model, and a one factor model in which all subtests loaded onto a single factor.
2. Multiple regression was used to test for variance explained on the WIAT-II) Word Reading and Numerical Operations subtests by the Shipley-2's composites that was beyond that explained by the WASI FSIQ. Semi-partial correlations were reported.

## Results

- Descriptive statistics for the sample's scores on the various measures are shown in Table 1.

Table1
*Descriptive statistics for Shipley-2 and WASI scores*

| Battery and Subtests | *N* | *M* | *SD* | Skewness | Kurtosis |
|---|---|---|---|---|---|
| **Shipley-2** | | | | | |
| Vocabulary | 100 | 80.19 | 21.80 | -0.69 | -0.19 |
| Abstraction | 100 | 78.30 | 20.63 | 0.05 | -0.83 |
| Block Patterns | 100 | 87.42 | 18.21 | 0.71 | -0.25 |
| Composite A | 100 | 76.27 | 22.71 | -0.27 | -0.55 |
| Composite B | 100 | 80.76 | 21.78 | -0.04 | -0.47 |
| **WASI** | | | | | |
| Vocabulary | 100 | 36.25 | 10.50 | 0.39 | -0.52 |
| Block Design | 100 | 43.53 | 11.34 | -0.04 | -0.45 |
| Similarities | 100 | 40.40 | 11.02 | -0.27 | -0.49 |
| Matrix Reasoning | 100 | 43.07 | 14.30 | -0.33 | -0.94 |
| Full Scale IQ | 100 | 86.84 | 16.49 | 0.32 | 0.83 |
| **WIAT-II** | | | | | |
| Vocabulary | 99 | 79.31 | 21.43 | -0.38 | -0.79 |
| Numerical- Operations | 99 | 84.05 | 18.91 | -0.53 | 1.58 |

Note. Shipley-2 and WIAT-II mean scores are in standard scores. WASI Full Scale IQ mean is a standard score, but WASI subtests are in T-Scores. Skewness and kurtosis for were in the acceptable ranges for maximum likelihood estimation (Curran, West, & Finch, 1996). All data were complete.

**CFA Results**

- CFA fit indices are presented in Table 2.

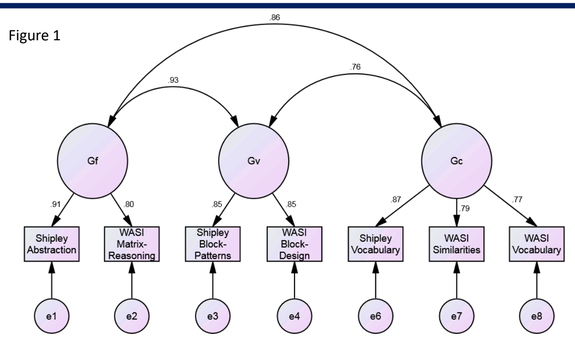
Table 2
*$\Delta\chi^2$ and fit statistics for three, two, and one factor models*

| Model | $\chi^2$ (*df*) | $\Delta\chi^2$ ($\Delta df$)* | RMSEA | CFI |
|---|---|---|---|---|
| Three-Factor (Gf, Gv, Gc) | 21.766 (11) | | .099 | .977 |
| Two-Factor (Gf/Gv, Gc) | 28.627 (13) | 6.861 (2) | .110 | .966 |
| One-Factor (Gf/Gv/Gc) | 57.154 (14) | 28.527 (1) | .176 | .907 |

Note. All changes in $\chi^2$ were statistically significant at the .05 level. *All comparisons are with the preceding model.

- Overall the three factor model (see figure 1) and the two factor model provide good fit, as indicated by CFI; the one factor model provided only adequate fit. There was a statistically significant $\Delta\chi^2$ between the three and two factor models, indicating that the three factor model provides the best fit.

Figure 1




**Multiple Regression Results**

- Results for the regression analysis are presented in Table 3.

Table 3
*Statistical significance values and semi-partial correlations for multiple regression analysis*

| Composite | *p* | Semi-partial correlation |
|---|---|---|
| **WIAT-II Word Reading** | | |
| Shipley-2 composite A | <.001 | .309 |
| WASI FSIQ | .095 | .110 |
| **WIAT-II Word Reading** | | |
| Shipley-2 composite B | <.001 | .292 |
| WASI FSIQ | <.01 | .185 |
| **WIAT-II Numerical Operations** | | |
| Shipley-2 composite A | .117 | .099 |
| WASI FSIQ | <.001 | .331 |
| **WIAT-II Numerical Operations** | | |
| Shipley-2 composite B | .415 | .052 |
| WASI FSIQ | <.001 | .426 |

- Analysis revealed that Shipley-2 composite A explained statistically significant variance in WIAT-II Word Reading scores above and beyond the variance explained by the WASI FSIQ. (i.e., the semi-partial correlation is the square root of the change in $R^2$.
- Both WASI FSIQ and Shipley-2 composite B explained statistically significant unique variance in WIAT-II Word Reading scores after the other composite's variance has been accounted for.
- WASI FSIQ explained statistically significant variance in Numerical Operations beyond that explained by Shipley-2 composite A
- The WASI FSIQ also explained statistically significant variance over the Shipley-2 composite B.

## Discussion

- The Shipley-2 manual indicates that the Shipley-2 measures two areas of cognition, Gf and Gc, but the results of our analysis indicate that it may measure three areas of cognition. Although the Shipley-2 Abstraction subtest appeared to fit well into a Gf factor, and the Shipley-2 Vocabulary subtest fit well into a Gc factor, the Shipley-2 Block Patterns subtest fit better into a separate Gv factor than into a combined Gf/Gv factor, according to statistical tests. It is important to mention that, from a practical standpoint, the Gf and Gv correlation was so strong that some may not consider these factors distinct.
  - If a practitioner is considering using either Block Patterns or Abstraction to measure a client's Gf abilities, Abstraction may be the better choice.
  - The Shipley-2's ability to measure an additional aspect of cognitive ability, spatial visualization, may be viewed as strength.
- Both the Shipley-2 and the WASI explained unique variance in academic achievement. These two batteries may be thought of as complimentary assessments rather than as interchangeable batteries when predicting academic achievement.
- Limitations include the below-average IQ of the sample used in analysis, and the limited number of subtests used in cross-battery CFA. It is also important to test the Shipley-2's ability to predict academic achievement across a broader array of achievement subtests. Future research with more robust batteries and individuals nearer to average IQ is needed.