Available online at www.sciencedirect.com

Intelligence 33 (2005) 417–430

INTELLIGENCE

Constructive replication of the visual–perceptual-image rotation model in Thurstone's (1941) battery of 60 tests of mental ability

Wendy Johnson*, Thomas J. Bouchard Jr.

Department of Psychology, University of Minnesota—Twin Cities, United States

Received 20 September 2004; received in revised form 20 November 2004; accepted 7 December 2004
Available online 12 February 2005

Abstract

We recently evaluated the relative statistical performance of the Cattell–Horn fluid–crystallized model and the Vernon verbal–perceptual model of the structure of human intelligence in a sample of 436 adults heterogeneous for age, place of origin, and educational background who completed 42 separate tests of mental ability from three test batteries. We concluded that the Vernon model's performance was substantively superior but could be significantly improved. In so doing, we proposed a four-stratum model with a *g* factor at the top of the hierarchy and three factors at the third stratum. We termed this the Verbal–Perceptual-Image Rotation (VPR) model. In this study, we constructively replicated the model comparisons and development of the VPR model using the data matrix published by Thurstone and Thurstone (1941) [Thurstone, L. L., & Thurstone T. G. (1941). *Factorial studies of intelligence*. Chicago: University of Chicago Press]. The data matrix was generated by scores of 710 Chicago eighth graders on 60 tests of mental ability.

Keywords: *g* factor; Fluid and crystallized intelligence; Verbal and perceptual abilities; Mental rotation; Spatial visualization; VPR theory

Scientific knowledge accumulates through the application of inductive inference. We articulate alternative hypotheses, carry out experiments with alternative outcomes and other studies designed to refute one or more of the hypotheses, and cycle back through this procedure by refining the possibilities

---

* Corresponding author. Department of Psychology, University of Minnesota, 75 East River Road, Minneapolis, MN 55455, USA. Tel.: +1 952 473 1673; fax: +1 952 473 1998.
E-mail address: john4350@umn.edu (W. Johnson).

0160-2896/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.intell.2004.12.001

that remain (Platt, 1964). Unlike deductive reasoning, this process of inductive inference relies upon the collection of empirical observations relevant to the proposition under evaluation, and we can never be certain that we have gathered all of the observations that may be relevant. In addition, the selection of which specific observations to make and the conditions in which to make them is always a matter of imagination, skill, and intuitive judgment. This means that the acceptance of empirical evidence for a proposition always involves some expression of confidence that the observations can be replicated, and renders actual replication of scientific studies crucial to the establishment of general acceptance for the proposition. This is especially true in a science like psychology, where the potential for truly experimental studies is extremely limited due to ethical constraints.

Lykken (1968) articulated three different methods of replication. Literal replication involves exact duplication of sampling procedure, experimental conditions, measuring techniques, and methods of analysis. This may be impossible to attain in practice, but it may not matter because the only real purpose of literal replication is to verify that the first study was conducted correctly. Operational replication involves duplication of only sampling and experimental procedures. Its purpose is to test whether the conditions and procedures described in the report of the original experiment produce the same result when operationalized by someone else. Constructive replication deliberately avoids duplication of procedures at all. Its purpose is to test the validity of the proposition independent of the methods used to test it, and it therefore provides the strongest and most general test of the proposition. The purpose of the current study was to constructively replicate our (Johnson & Bouchard, 2005) observation that Vernon's (1964, 1965) verbal–perceptual model of the structure of human intelligence offers a better description of the structure of human intelligence than the Cattell–Horn (Cattell, 1971; Horn, 1989) fluid–crystallized model. In addition, we constructively replicated the development of our verbal–perceptual-image rotation (VPR) model, which appears to provide significantly better statistical fit than either of the other two.

The Cattell–Horn (Cattell, 1971; Horn, 1989) fluid–crystallized theory of the structure of human intellect proposes an important distinction between fluid and crystallized ability factors. The theory has tended to dominate discussions of the structure of intellect in the past 40 years. Fluid ability reflects the capacity to solve problems for which prior experience and learned knowledge and skills are of little use. Crystallized ability reflects consolidated knowledge gained by education, access to cultural information, and experience. Following their conceptualization, we (Johnson & Bouchard, 2005) implemented the fluid–crystallized model with no general intelligence, or *g*, factor but as a hierarchy of three strata with broad fluid and crystallized ability factors at the top, six more specific ability factors (fluid, crystallized, visualization, memory, apprehension speed, and facility) in the middle stratum, and the administered tests in the first stratum.

Though the fluid–crystallized model has tended to dominate discussion, other models have been suggested. One of the most prominent of these was proposed by Vernon (1964, 1965). He stressed the importance of *g* in contributing to all mental abilities, but observed that, below *g* in the hierarchy, the abilities tend to fall in two main groups that refer to verbal and educational abilities, and to spatial, practical, and mechanical abilities. With the intention of following Vernon's conceptualization while making only theoretically important distinctions from the fluid–crystallized model, we (Johnson & Bouchard, 2005) implemented his model as a hierarchy with four strata: *g* at the top, verbal and perceptual ability factors in the third stratum, six more specific ability factors (verbal, scholastic, fluency, number, perceptual speed, and spatial) in the second stratum, and the administered tests in the first stratum. We included no memory factor as Vernon suggested that memory demands tend to be distributed relatively evenly over the other abilities (Vernon, 1964). Vernon also viewed a reasoning factor as generally unnecessary, as he believed that reasoning ability is generally reflected in *g*.

Our original observations (Johnson & Bouchard, 2005) were made in a sample of 436 heterogeneous for age, place of origin, and educational background who completed 42 separate tests of mental ability from three test batteries. In that sample, the Vernon verbal–perceptual model fit the data substantially better than did the Cattell–Horn fluid–crystallized model. (We also compared a model based on Carroll's (1993) conceptualization of the structure of intellect in that sample, but the results were so similar to those of the fluid–crystallized model that we have not included this model in the current replication study.) At the same time, neither model fit the data well according to the fit criteria we had established. By adding second stratum memory and image rotation factors as well as an image rotation factor at the third stratum, we were able to improve the fit significantly and easily meet the fit criteria we had established. It is this model we termed VPR. The potential to replicate these findings constructively in other samples and combinations of tests is complicated by the dependence of the content of the factors that can be identified in any factor model on the specific content of the tests contained in the battery, as well as the potential for demographic characteristics of the sample to produce factors of their own in the data. This emphasizes that successful constructive replication offers greater evidence in support of any given set of findings when the samples and contents of the test batteries are divergent from those in the original study.

1. Method

1.1. Sample

For this study, we made use of the data matrix of 60 mental ability tests published by Thurstone and Thurstone (1941) in their *Factorial Studies of Intelligence*. It was clear that a large battery of tests was necessary in order to generate sufficient detail to replicate our original observations (Johnson & Bouchard, 2005); at the same time, the larger the battery, the stronger the test, as there was less likelihood that we would merely reproduce commonly agreed-upon content factors. The sample on which the matrix was based consisted of 710 Chicago area eighth graders, reasonably evenly distributed among 15 different public schools. The tests were administered in eleven 1-h sessions over a period of about 2 weeks. The authors did not specify the numbers of males and females or the distribution of ages. Nevertheless, the sample must have been relatively homogeneous, particularly in comparison to the sample on which our original observations were based, due to the similarity in age and area of residence.

1.2. Measures

The tests used by Thurstone and Thurstone (1941) were adapted for use with younger children from others used by the authors with high school and college students. In addition, several new tests were selected from previous experimental batteries. The 60 tests are summarized in Table 1. Though their content spanned the same general range of ability as the 42 tests on which our original observations (Johnson & Bouchard, 2005) were based and there was some reasonably close overlap, there were also substantial differences. Specifically, both batteries included the test Pedigrees and both had

Table 1
Tests included in the battery

Test	Assessment activity
1. ABC	Given letter combination codes, decode letter sequences.
2. Absurdities	Read sentences and label them sensible or foolish.
3. Addition	Check whether columns of addition are correct.
4. Anagrams	Given a word, build anagrams of its letters.
5. Arithmetic	Complete arithmetic word problems and select the correct answer from 5 options.
6. Association	Write words that are names of things to eat or drink.
7. Mirror reading	Identify the reverse-printed version of a word from 4 options.
8. Cards	Identify 2-D rotated versions of presented figures.
9. Classification	Classify sets of numbers in assigned categories.
10. Completion	Think of a word to complete a given sentence and select its first letter from 5 options.
11. Digit span	Write sequences of 4 digits upon recital by examiner.
12. Directions	Follow written instructions presented in brief sentences.
13. Disarranged sentence	Rearrange presented words to make sensible sentences.
14. Dot counting I	Count the numbers of dots in presented rows.
15. Dot counting II	Count the numbers of dots in presented squares.
16. Dot counting III	Count the numbers of dots in presented patterns.
17. Dot patterns	Identify dot patterns that contain different numbers of dots.
18. Faces	Identify the face that is different from 3 presented in a row.
19. Figure grouping	Identify the figure that is different from 5 presented in a row.
20. Figure naming	Write the first letters of the names of presented figures.
21. Figure recognition	Study presented figures and identify them when placed with others.
22. Figures	Distinguish presented figure from mirror-image versions when 2-D rotated.
23. First and last letters	Write words having specified first and last letters.
24. First letters	Write words having specified first letters.
25. First names	Study a list of first and last names and write the first name when presented with the last.
26. Flags	Distinguish presented flags from mirror-image versions when 2-D rotated.
27. Four-letter words	Write four-letter words beginning with specified letters.
28. Geometrical forms	Place an x in a collection of figures according to specific instructions.
29. High numbers	Identify the numbers higher than the adjoining numbers in a row of numbers.
30. Identical numbers	Identify the numbers in a column identical to the one at the top.
31. Identical pictures	Identify pictures in a row identical to the first.
32. Incomplete words	Identify words with missing letters.
33. Letter grouping	Identify the four-letter series that has a different pattern from a group of 4.
34. Letter series	Specify the next letter in a series.
35. Mazes I	Draw a line from point A to point B through a maze.
36. Mazes II	Draw a line from point A to point B through a maze.
37. Multiplication	Check whether multiplication problems are correct.
38. Number patterns	Specify the number left out of a 4×4 square of numbers.
39. Paragraph recall	Listen to examiner read a paragraph and fill in the missing words in a printed version.
40. Pedigrees	Answer questions regarding relationships among members of a pedigree tree.
41. Picture naming	Write the first letters of the names of presented pictures.
42. Prefixes	Write words having specified prefixes.
43. Proverbs	Identify the proverb with a different meaning from a list of 5.
44. Pursuit	Follow certain lines from top of figure to bottom in a figure with many intersecting lines.
45. Reading vocabulary	Identify word synonymous with given word.
46. Sentence completion	Choose the word appropriate to complete a given sentence.
47. Reading comprehension I	Read paragraphs and answer questions about them.
48. Reading comprehension II	Read paragraphs and answer questions about them.

Table 1 (*continued*)

Test	Assessment activity
49. Reasoning	Given two relationships, specify a third.
50. Rhyming words	Write four words that rhyme with the given word.
51. Same or opposite	Identify the word that is either the same or the opposite of the given word from a list of 4.
52. Scattered *X*'s	Circle the *x*'s in a random scattering of letters.
53. Secret writing	Given three words and a coding for them, re-identify the words.
54. Suffixes	Write words ending with specified suffixes.
55. Synonyms	Write three words synonymous with the given word.
56. Three-higher	Identify numbers three higher than the previous number in a row of 12.
57. Verbal enumeration	Identify words in a column that fall into the category labeled at the top.
58. Word checking	Identify all words from a list that meet the given specifications.
59. Word–Number recall	Study pairs of objects and numbers, then recall the numbers when given the objects.
60. Word puzzles	Unscramble letters to produce words falling in given categories.

vocabulary, arithmetic computation, identical pictures, digit span and proverb interpretation tests. On the other hand, our original battery included several tests requiring three-dimensional mental rotation and Thurstone's battery included none. Our original battery also included the Wechsler Adult Intelligence Scale, which primarily requires free responses, while Thurstone's battery was almost completely multiple choice, with the exception of some of the fluency tests. Thurstone's battery also included several dot counting tasks that were very different from anything in our original battery, as well as many more identification tasks.

1.3. Statistical analyses

We used the same statistical procedures to develop the models for this study as we had used in making our original observations (Johnson & Bouchard, 2005). Thus we used maximum likelihood confirmatory factor analysis to implement the models we evaluated. As we described previously, because the standard and most readily available chi-square measure of model fit generally indicates significant lack of fit in large samples, we used a chi-square statistic less than 2* degrees of freedom and Root Mean Square Error of Approximation (RMSEA) less than .05 (Browne & Cudeck, 1992) as indications of good model fit. RMSEA measures the extent of discrepancy between the model and data per degree of freedom; thus both alternative fit statistics provide some benefit to more parsimonious models. Model comparison among non-nested models by necessity relies upon information-theoretic fit statistics that emphasize minimization of the amount of information required to express the modeled data, which results in the selection of models that are the most parsimonious or efficient representations of the observed data. To compare our models, we made use of one of these fit statistics, the Bayesian Information Criterion (BIC; Raftery, 1995). BIC estimates the natural log of the ratio of the posterior to the prior odds of the model in comparison to its saturated version, favoring the saturated model when positive and the presented model when negative. A difference in BIC of 10 between non-nested models is considered clear evidence in favor of the model with the more negative BIC. BIC heavily favors more parsimonious models.

In implementing the Cattell–Horn fluid–crystallized and Vernon verbal–perceptual models for comparison, we allowed nine second-stratum factors in each in order to ensure that the models were as directly comparable as possible. We chose the number nine because the resulting factors could be justified within each of the theoretical frameworks, allowed each of the tests (with the exception of Word–Number Recall) to load reasonably on at least one of the nine, and developed substantial loadings from several to many tests. In addition, the data correlation matrix had nine eigenvalues greater than 1.0, giving the use of nine factors an objective and generally accepted rationale. We note that Thurstone and Thurstone (1941) presented a solution with 10 factors, but the last three were essentially doublets. To maximize consistency of results, we assigned first-stratum tests to load on second-stratum factors uniformly unless there were substantive theoretical reasons or reasons of factor content to assign them differently. In both models, we allowed three sets of residual correlations resulting from analysis of modification indexes and consideration of test content. These residual correlations were between Word–Number Recall and First Names (which both require recall of otherwise unassociated pairs when presented with one member of the pair), between Mazes I and II (which have otherwise unique highly similar content), and between Prefixes and Suffixes (which also have highly similar content otherwise unique).

To replicate the development of the VPR model in this sample, we followed the same procedures we had used in the development of the model in our original sample (Johnson & Bouchard, 2005). That is, we added to the verbal–perceptual model a second-stratum content memory factor and a factor capturing the tests involved in negative cross-loadings at both second- and third-strata.

2. Results

For the fluid–crystallized model (Fig. 1), we obtained a chi-square of 6501.82 on 1654 *df*, $p<0.00001$, with RMSEA=0.064 and BIC=−4357.13. The correlation between the two third-stratum factors, representing original fluid and crystallized intelligence, was 0.99, indicating that, in this battery, fluid

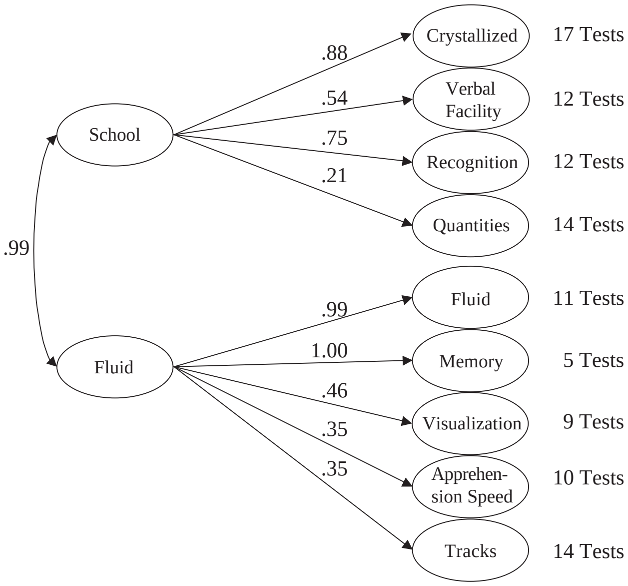

Fig. 1. Structural portion of fluid–crystallized model. Chi-squre=6501.82, *df*=1654, RMSEA=0.064, BIC=−4357.13. Some tests have primary and secondary factor loadings.

and crystallized intelligence were effectively equivalent. In addition, the loading of the second-stratum fluid factor on the third-stratum original fluid factor was 0.99, the loading of the second-stratum memory factor on the third-stratum original fluid factor was 1.00, and the correlation between the 2 third-stratum factors was 0.99. This suggested that these two second-stratum factors were indistinguishable both from each other and from their third-stratum counterpart.

For the verbal–perceptual model (Fig. 2), we obtained a chi-square of 4608.86 on 1659 *df*, $p<0.00001$, with RMSEA=0.050 and BIC=−6282.91. The correlation between the *v*:*ed* and *k*:*m* factors was 0.75. This clearly defined a general intelligence factor, though we constrained the loadings of the *v*:*ed* and *k*:*m* factors on the general factor equal in order to identify the model. The correlation between these two factors in this sample was almost identical to that from our original observations (Johnson & Bouchard, 2005; 0.76). The highest loading of a second-stratum factor on its third-stratum factor was 0.91, so each of the second-stratum factors was distinguishable from the third-stratum factors.

As with our original observations (Johnson & Bouchard, 2005), neither model fit really well by the fit criteria we had established, though both performed reasonably. Again, the verbal–perceptual model fit substantially better than did the fluid–crystallized model, as indicated by both RMSEA and BIC. The difference in BIC of about 1926 suggests that the odds of the verbal–perceptual model being correct over the fluid–crystallized model are in excess of 10^{100} to 1, rendering their actual quantification unimportant.

In addition to the model-fitting results, evaluation of the loadings from the first to the second strata is also relevant to assessment of the appropriateness of the two models. These loadings are shown in Table 2, and the unique variable and latent factor variances are shown in Table 3. As the table indicates, there were many tests that we expected to show substantive loadings on the crystallized factor in that model, as they draw on factual knowledge. Yet three of these tests (Figure Names, Incomplete Words, and Picture Names) showed significant *negative* loadings on the crystallized factor. In addition, three of the dot-counting tests showed significant negative loadings on the visualization factor. The verbal–perceptual model did not generate anomalous results of this nature, though the absence of a memory

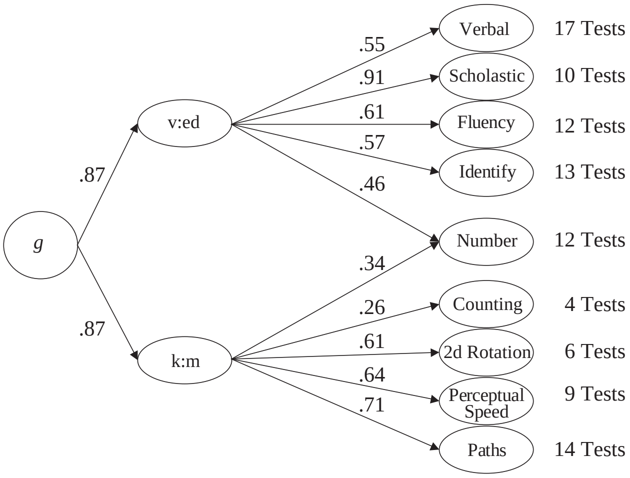

Fig. 2. Structural portion of verbal–perceptual model. Chi-squre=4608.86, *df*=1659, RMSEA=0.050, BIC=−6282.91. Some tests have primary and secondary factor loadings.

Table 2
Loadings from first to second strata for the fluid–crystallized, verbal–perceptual, and VPR models

Test	Primary loadings						Secondary loadings					
	Fluid–crystallized		Verbal–perceptual		VPR		Fluid–crystallized		Verbal–perceptual		VPR	
1. ABC	fluid	0.37	number	0.31	scholastic	0.29	quantities	0.32	scholastic	0.28	number	0.28
2. Absurdities	fluid	0.52	verbal	0.48	verbal	0.46	–		paths	0.14	paths	0.17
3. Addition	quantities	0.58	number	0.53	number	0.56	crystallized	0.19	–		speed of rot.	0.44
4. Anagrams	verbal facility	0.55	fluency	0.54	fluency	0.56	appre. speed	0.30	perc. speed	0.27	perc. speed	0.27
5. Arithmetic	crystallized	0.54	number	0.41	number	0.46	quantities	0.18	verbal	0.36	verbal	0.31
6. Association	verbal facility	0.48	fluency	0.54	fluency	0.53	recognition	0.32	identify	0.27	identify	0.28
7. Mirror reading[a]	appre. speed	0.49	perc. speed	0.51	perc. speed	0.47	verbal facility	0.32	fluency	0.29	fluency	0.32
8. Cards	visualization	0.82	2d rotation	0.85	2d rotation	0.84	–		–		–	
9. Classification	fluid	0.39	number	0.33	number	0.35	quantities	0.19	scholastic	0.23	scholastic	0.21
10. Completion[b]	crystallized	0.62	verbal	0.75	verbal	0.67	recognition	0.27	identify	0.23	identify	0.18
11. Digit span	memory	0.31	scholastic	0.36	cont. memory	0.36	–		–		–	
12. Directions[c]	crystallized	0.45	verbal	0.47	verbal	0.41	recognition	0.28	number	0.25	number	0.28
13. Disarranged sentences	recognition	0.55	identify	0.44	identify	0.48	crystallized	0.19	verbal	0.43	verbal	0.38
14. Dot counting I[d]	quantities	0.69	counting	0.74	counting	0.63	tracks	0.27	number	0.23	speed of rot.	0.56
15. Dot counting II[d]	tracks	0.55	counting	0.52	counting	0.44	quantities	0.47	paths	0.43	paths	0.44
16. Dot counting III[d]	quantities	0.59	counting	0.83	counting	0.80	tracks	0.40	paths	0.18	speed of rot.	0.53
17. Dot patterns[d]	quantities	0.54	number	0.38	paths	0.38	tracks	0.35	counting	0.28	speed of rot.	0.35
18. Faces	tracks	0.56	paths	0.53	paths	0.57	appre. speed	0.37	perc. speed	0.31	perc. speed	0.27
19. Figure grouping	tracks	0.49	paths	0.53	paths	0.53	–		–		–	
20. Figure naming	recognition	0.68	identify	0.55	identify	0.53	crystallized	−0.34	–		speed of rot.	0.21
21. Figure recognition	memory	0.28	scholastic	0.32	scholastic	0.31	–		–		–	
22. Figures	visualization	0.72	2d rotation	0.71	2d rotation	0.71	–		–		–	
23. First and last letters	verbal facility	0.68	fluency	0.68	fluency	0.68	–		–		–	
24. First letters	verbal facility	0.78	fluency	0.78	fluency	0.78	–	–	–			
25. First names	memory	0.40	scholastic	0.46	cont. memory	0.45	–		–		–	
26. Flags	visualization	0.74	2d rotation	0.74	2d rotation	0.74	–		–		–	
27. Four-letter words	verbal facility	0.66	fluency	0.65	fluency	0.65	–		–		–	
28. Geometrical forms	visualization	0.34	paths	0.32	paths	0.34	tracks	0.32	2d rotation	0.30	2d rotation	0.28
29. High numbers[e]	tracks	0.35	number	0.44	number	0.44	quantities	0.26	paths	0.31	paths	0.33
30. Identical numbers[f]	appre. speed	0.40	identify	0.38	speed of rot.	0.44	quantities	0.33	perc. speed	0.36	perc. speed	0.39
31. Identical pictures	tracks	0.51	paths	0.52	paths	0.52	appre. speed	0.35	perc. speed	0.24	perc. speed	0.25
32. Incomplete words[g]	verbal facility	0.41	fluency	0.39	fluency	0.40	recognition	0.37	identify	0.29	identify	0.30
33. Letter grouping	fluid	0.47	scholastic	0.62	scholastic	0.62	–		–		–	
34. Letter series	fluid	0.59	scholastic	0.70	scholastic	0.71	–		–		–	
35. Mazes I	tracks	0.57	paths	0.57	paths	0.57	–		–		–	

36. Mazes II	tracks	0.39	paths	0.37	paths	0.38	visualization	0.20	2d rotation	0.20	2d rotation	0.19
37. Multiplication[h]	quantities	0.61	number	0.42	speed of rot.	0.53	crystallized	0.18	identify	0.33	number	0.48
38. Number patterns	appre. speed	0.35	perc. speed	0.31	perc. speed	0.29	quantities	0.19	number	0.22	number	0.26
39. Paragraph recall	memory	0.68	verbal	0.73	verbal	0.66	–		–		cont. memory	0.12
40. Pedigrees	fluid	0.65	scholastic	0.49	scholastic	0.52	–		verbal	0.26	verbal	0.24
41. Picture naming	identify	0.92	identify	0.67	identify	0.65	crystallized	−0.51	–		speed of rot.	0.17
42. Prefixes	verbal facility	0.64	fluency	0.64	fluency	0.64	–		–		–	
43. Proverbs	crystallized	0.53	verbal	0.61	verbal	0.61	fluid	0.11	–		–	
44. Pursuit	tracks	0.70	paths	0.67	paths	0.66	–		–		–	
45. Reading vocabulary	crystallized	0.87	verbal	0.88	verbal	0.89	–	–	–			
46. Sentence completion	crystallized	0.84	verbal	0.85	verbal	0.86	–		–		–	
47. Reading comprehension I	fluid	0.67	verbal	0.75	verbal	0.76	–		–		–	
48. Reading comprehensino II	fluid	0.59	verbal	0.65	verbal	0.65	–		–		–	
49. Reasoning	fluid	0.46	number	0.30	number	0.34	–		verbal	0.26	verbal	0.22
50. Rhyming words[i]	verbal facility	0.59	fluency	0.59	fluency	0.58	recognition	0.26	identify	0.16	identify	0.17
51. Same or opposite	crystallized	0.68	verbal	0.75	verbal	0.75	recognition	0.16	identify	0.14	identify	0.15
52. Scattered *X*'s	appre. speed	0.33	perc. speed	0.35	perc. speed	0.38	–		–		speed of rot.	0.07
53. Secret writing	fluid	0.40	scholastic	0.55	scholastic	0.57	appre. speed	0.31	2d rotation	0.20	2d rotation	0.17
54. Suffixes	verbal facility	0.68	fluency	0.67	fluency	0.67	–		–		–	
55. Synonyms	verbal facility	0.58	fluecny	0.57	fluency	0.57	–		–		–	
56. Three-higher	quantities	0.46	number	0.76	number	0.76	crystallized	0.40	–		–	
57. Verbal enumeration	recognition	0.76	identify	0.59	identify	0.63	crystallized	0.04	verbal	0.39	verbal	0.33
58.Word checking	recognition	0.46	verbal	0.46	verbal	0.41	crystallized	0.25	identify	0.35	identify	0.40
59. Word–number recall	memory	0.05	scholastic	0.17	cont. memory	0.16	–		–		–	
60. Word puzzles	verbal facility	0.60	fluency	0.59	fluency	0.62	appre. speed	0.39	perc. speed	0.36	perc. speed	0.36

VPR model refers to VerbalPerceptual-Image Rotation model.

[a] Mirror reading had a tertiary loading of 0.26 on tracks in the fluid-crystallized model and 0.18 on paths in the verbal–perceptual model. It had a loading of 0.21 on paths in the VPR model.

[b] Completion had a tertiary loading of 0.14 on content memory in the VPR model.

[c] Directions had a tertiary loading of 0.16 on number in the fluid-crystallized model and 0.22 on identify in both the verbal–perceptual and VPR models.

[d] All the Dots tests had tertiary loadings in the fluid-crystallized model on visualization. For Dots I–III, the loadings were negative, ranging from −0.26 to −0.18. For Dot patterns, the loading was 0.06. In the verbal–perceptual model. Dot patterns had a tertiary loading of 0.28 on paths. In the VPR model, Dots I had a tertiary loading of 0.21 on number, Dots II had a tertiary loading of 0.41 on 3d rotation. Dots III had a tertiary loading of 0.11 on paths; and Dot patterns had tertiary loadings of 0.33 on number and 0.18 on counting.

[e] High numbers had a tertiary loading in the fluid-crystallized model of 0.23 on crystallized.

[f] Identical numbers had subsidiary loadings in the fluid-crystallized model of 0.18 on tracks and 0.10 on recognize. It had a tertiary loading of 0.13 on paths in the verbal–perceptual model. In the VPR model, it had tertiary loadings of 0.28 on identify and 0.27 on paths.

[g] Incomplete words had subsidiary loadings in the fluid-crystallized model of 0.26 on apprehension speed and −0.16 on crystallized. It had a tertiary loading of 0.21 on perceptual speed in the verbal–perceptual model and 0.22 in the VPR model.

[h] Multiplication had a tertiary loading of 0.25 on perceptual speed in the VPR model.

[i] Rhyming words had a tertiary loading of 0.15 on verbal–perceptual model and 0.14 on verbal in the VPR model.

Table 3
Unique variable and latent factor variances from the Cattell–Horn fluid–crystallized, Vernon verbal–perceptual, and VPR models

Variable	Fluid–crystallized	Verbal–perceptual	VPR
1. ABC	0.71	0.71	0.72
2. Absurdities	0.72	0.71	0.71
3. Addition	0.59	0.72	0.56
4. Anagrams	0.55	0.55	0.55
5. Arithmetic	0.64	0.59	0.58
6. Association	0.55	0.54	0.54
7. Mirror reading	0.48	0.42	0.44
8. Cards	0.32	0.28	0.29
9. Classification	0.79	0.74	0.73
10. Completion	0.27	0.28	0.28
11. Digit span	0.90	0.87	0.87
12. Directions	0.49	0.48	0.48
13. Disarranged sentences	0.53	0.51	0.51
14. Dot counting I	0.42	0.33	0.34
15. Dot counting II	0.46	0.46	0.47
16. Dot counting III	0.45	0.22	0.20
17. Dot patterns	0.54	0.53	0.53
18. Faces	0.50	0.48	0.49
19. Figure grouping	0.76	0.72	0.73
20. Figure naming	0.72	0.70	0.70
21. Figure recognition	0.91	0.90	0.90
22. Figures	0.48	0.49	0.50
23. First and last letters	0.53	0.53	0.53
24. First letters	0.39	0.39	0.39
25. First names	0.83	0.79	0.80
26. Flags	0.46	0.45	0.45
27. Four-letter words	0.57	0.57	0.57
28. Geometrical forms	0.75	0.72	0.71
29. High numbers	0.67	0.58	0.57
30. Identical numbers	0.60	0.57	0.47
31. Identical pictures	0.57	0.56	0.56
32. Incomplete words	0.56	0.54	0.54
33. Letter grouping	0.78	0.61	0.61
34. Letter series	0.65	0.51	0.50
35. Mazes I	0.67	0.67	0.68
36. Mazes II	0.79	0.76	0.74
37. Multiplication	0.55	0.60	0.41
38. Number patterns	0.83	0.80	0.81
39. Paragraph recall	0.52	0.47	0.47
40. Pedigrees	0.58	0.57	0.56
41. Picture naming	0.51	0.55	0.57
42. Prefixes	0.59	0.59	0.59
43. Proverbs	0.61	0.62	0.63
44. Pursuit	0.51	0.55	0.57
45. Reading vocabulary	0.25	0.22	0.21
46. Sentence completion	0.30	0.27	0.27
47. Reading comprehension I	0.55	0.44	0.43

Table 3 (*continued*)

Variable	Fluid–crystallized	Verbal–perceptual	VPR
48. Reading comprehension II	0.65	0.58	0.58
49. Reasoning	0.79	0.78	0.78
50. Rhyming words	0.46	0.47	0.47
51. Same or opposite	0.37	0.35	0.35
52. Scattered *X*'s	0.89	0.88	0.86
53. Secret writing	0.67	0.57	0.56
54. Suffixes	0.54	0.55	0.55
55. Synonyms	0.67	0.67	0.67
56. Three-higher	0.57	0.42	0.43
57. Verbal enumeration	0.38	0.35	0.35
58. Word checking	0.89	0.88	0.56
59. Word–number recall	0.67	0.57	0.98
60. Word puzzles	0.54	0.55	0.40
Latent variables from fluid–crystallized model			
Fluid intelligence	0.02	n/a	n/a
Crystallized intelligence	0.23	n/a	n/a
Quantities	0.95	n/a	n/a
Verbal facility	0.71	n/a	n/a
Apprehension speed	0.88	n/a	n/a
Visualization	0.79	n/a	n/a
Memory	0.01	n/a	n/a
Recognize	0.44	n/a	n/a
Tracks	0.88	n/a	n/a
Latent variables from verbal–perceptual models			
Verbal	n/a	0.70	0.72
Scholastic	n/a	0.17	0.19
Fluency	n/a	0.63	0.64
Number	n/a	0.44	0.39
Content memory	n/a	n/a	0.01
Perceptual speed	n/a	0.59	0.75
Counting	n/a	0.93	0.73
Identify	n/a	0.67	0.60
Paths	n/a	0.49	0.36
2d rotation	n/a	0.63	0.55
Speed of rotation	n/a	n/a	0.00

For the fluid–crystallized model, there were unique variable variance correlations of 0.27 between Word–Number Recall and First Names, 0.42 between Mazes I and II, and 0.12 between Prefixes and Suffixes. For the verbal–perceptual model, the analogous correlations were 0.21, 0.42, and 0.12. For the VPR model, the analogous correlations were 0.23, 0.41, and 0.12.

factor appeared to be important in this sample as it was in our original sample (Johnson & Bouchard, 2005). In particular, the Word–Number Recall test did not generate substantial loadings on any factor, though we note that this test did not generate a substantial loading on the memory factor in the fluid–crystallized model either.

Implementation of the VPR model (Fig. 3 and Tables 2 and 3) resulted in a chi-square of 4055.94 on 1646 *df*, with RMSEA=0.045 and BIC=−6750.49. Again, the VPR model offered a significant and

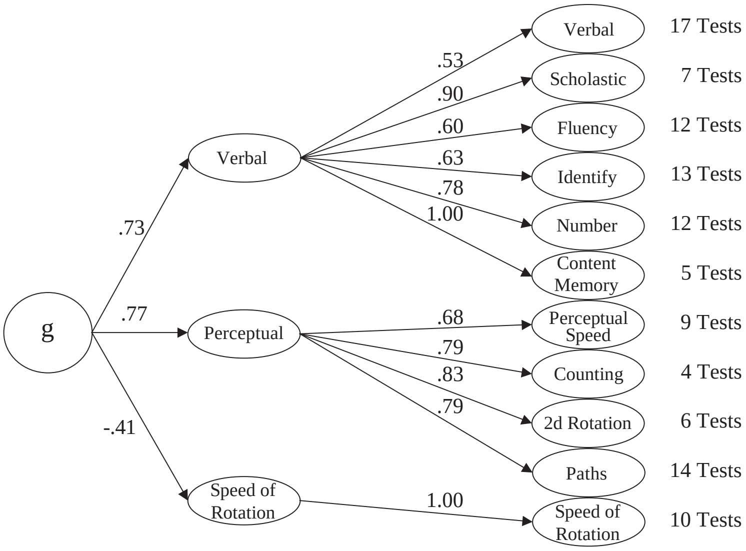

Fig. 3. Structural portion of verbal–perceptual-image rotation model. Chi-squre=4055.94, *df*=1646, RMSEA=0.045, BIC=−6750.49. Some tests have primary and secondary factor loadings.

substantive improvement in fit, with the odds of the VPR model being correct over the verbal–perceptual model being in excess of 10^{100} to 1. We note that this model still did not meet the specified fit criterion of having a chi-square statistic less than 2* degrees of freedom. It did, however, fit comparably to that from our original sample of 436 (Johnson & Bouchard, 2005) in the sense that, when the sample size was adjusted from 710 to 436, the chi-square was 2488.48 on 1646 *df* (RMSEA=0.034), which did meet the specified fit criterion. Interestingly, in these data, the second-stratum memory factor (termed content memory to distinguish it from the memory factor in the fluid–crystallized model) loaded on the third-stratum verbal factor rather than the third-stratum perceptual factor, as was the case in our original observations (Johnson & Bouchard, 2005). In addition, though the correlation between the third-stratum verbal and perceptual factors was strong and positive (0.72), the correlations of each of these factors with the image rotation factor were negative (−0.17 with verbal and −0.38 with perceptual).

3. Discussion

The primary goal of this study was to replicate constructively our findings (Johnson & Bouchard, 2005) (1) that Vernon's verbal–perceptual model offers a better description of the structure of intellect than does Cattell and Horn's fluid–crystallized model and (2) that substantial improvement in descriptive power can be gained by explicitly recognizing mental image rotation ability at a level comparable to that of verbal and perceptual ability. In spite of the use of two very demographically different samples taking two very different batteries of tests, our current results were very consistent with our earlier results. Thus, this study provides important additional evidence in support of the conclusion that there is a general intelligence factor contributing substantively to all aspects of intelligence. It also provides additional evidence that a distinction between verbal and perceptual abilities is more appropriate than the pervasive distinction between fluid and crystallized intelligence in describing the structure of intellect. Finally, it adds to the substantial body of evidence that spatial visualization/image rotation processes have not been given the attention they deserve as important and relatively independent contributors to the manifestation of human intelligence.

Though our results here, based on Thurstone and Thurstone's (1941) data, were very similar to our previous results (Johnson & Bouchard, 2005), there were two differences in the implemented VPR model worthy of note. First, the second-stratum memory factor loaded on the third-stratum verbal factor here rather than on the third-stratum perceptual factor as in our previous data. We interpreted this as being consistent with Vernon's (1964, 1965) description of memory as tending to be distributed across verbal and perceptual domains, depending on the nature of the memory requirements in any given task. The memory tasks in Thurstone and Thurstone's battery all directly involved verbal recall, while several of those in the battery on which our previous results were based were explicitly perceptual in nature.

Second, though the relationships between the third-stratum verbal and perceptual factors in the VPR models in the two studies were similar (correlations of 0.73 here and 0.81 in our previous results), the relationships between the third-stratum image rotation factors in the two studies and the other two third-stratum factors (verbal and perceptual) were not. In particular, the correlations here were both negative (−0.17 with verbal and −0.37 with perceptual), but they were positive in our previous results (0.42 with verbal and 0.85 with perceptual). We interpreted the factor here as image rotation speed. Thurstone and Thurstone's battery contained no tests that explicitly required three-dimensional mental image rotation, but the tests that loaded on the image rotation factor all involved speeded perception in circumstances in which mental image rotation might offer an effective strategy (Addition, Dots I–III and Dot Patterns, Figure Naming, Identical Numbers, Multiplication, Picture Naming, and Scattered *X*'s). There is substantial evidence that people use different strategies in solving spatial problems (Hegarty & Waller, 2004; Schultz, 1991), and that people use three-dimensional mental image rotation in some situations in which it is not explicitly required (Tarr, 1995; Tarr & Pinker, 1990). The battery on which our previous results were based contained four tests that did explicitly require three-dimensional mental image rotation, and these were the tests that loaded on the visualization factor in that data. We speculate that the use of three-dimensional mental image rotation is the link between the third-stratum image rotation factors in the two studies, but this needs to be investigated further in future research.

The results of these two studies seem conclusive with respect to the relative descriptive power of the fluid–crystallized and Vernon models. The Vernon model offers a superior description of the structure of human intellectual abilities. We believe that this is because the fluid–crystallized model refers to the process by which specific skills are developed and maintained rather than to the underlying structure of ability that makes acquisition of the skills possible. This has important implications for the study of brain structure and function. The refinements to the Vernon model we made to propose the VPR model must of necessity be considered open-ended by comparison. All models of this type are dependent to some degree on the specific tests included, with the range of abilities sampled as well as the level of specificity with which they are sampled determining the ability to distinguish among them (Horn, 1989). Still, our current study strongly corroborates our earlier conclusion that, at a high level of generalization, spatial visualization/image rotation abilities can be distinguished from verbal and perceptual abilities, independently of g. Along with g, all three of these types of abilities contribute importantly to the execution of specific human ability tasks.

References

Browne, M., & Cudeck, R. (1992). Alternative methods of assessing model fit. *Sociological Methods & Research*, *21*, 230–258.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, England: Cambridge University Press.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton-Mifflin.

Hegarty, M., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*, 175–191.

Horn, J. L. (1989). Models of intelligence. In R. L. Linn (Ed.), *Intelligence: measurement, theory, and public policy* (pp. 29–73). Urbana, IL: University of Illinois Press.

Johnson, W., & Bouchard Jr., T. (2005). Testing the grand old models of the structure of human intelligence: It's verbal, perceptual, and visualization (VPZ), not fluid and crystallized. *Intelligence*.

Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin*, *70*, 151–159.

Platt, J. R. (1964). Strong inference. *Science*, *146*, 347–353.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163.

Schultz, K. (1991). The contribution of solution strategy to spatial performance. *Canadian Journal of Psychology*, *45*, 474–491.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin and Review*, *2*, 55–82.

Tarr, M. J., & Pinker, S. (1990). When does human object recognition use a viewer-centered reference frame? *Psychological Science*, *1*, 253–256.

Thurstone, L. L., & Thurstone, T. G. (1941). *Factorial studies of intelligence*. Chicago: University of Chicago Press.

Vernon, P. (1964). *The structure of human abilities*. London: Muthen and Co. Ltd.

Vernon, P. (1965). *Ability factors and environmental influences*. *American Psychologist*, *20*, 723–733.