Slide 1

BAYLOR UNIVERSITY
SCHOOL OF EDUCATION

Are we Over-*Interpreting Students' Performance on Tests of Intelligence?*

A Re-Analysis of the Foundations of CHC Theory

Nicholas F. Benson
Alexander A. Beaujean
Ashley Donohue
Hailin Chi

Slide 2

Agenda

- Intro to the Cattell-Horn-Carroll (CHC) theory and its foundations
- Need for Study
- Our method and results
- Theoretical implications
- Practical implications

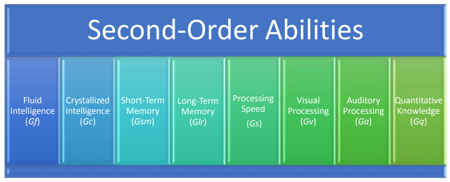
Slide 3

Introduction

CHC Theory and its Foundations

Slide 4

Cattell-Horn *Gf-Gc* Theory



Intelligence represents effects and interactions of numerous abilities working in concert. *Gf* and *Gc* viewed as more general abilities that support the others, *g* is not in the model.
.

Slide 5

Carroll's Three-Stratum Theory

- Strata distinguished by generality (breadth) and abstraction of abilities
- Direct hierarchical (bifactor) structure (Beaujean, 2015)
  - *g* and group factors have direct effects on measured abilities
  - *g* and group factors are orthogonal
- Provides the corpus of evidence for CHC theory
  - Frequently cited as empirical basis for interpreting lower strata abilities

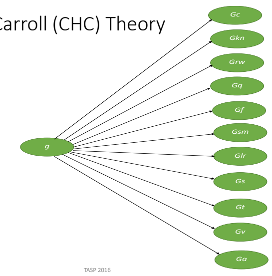
Slide 6

Cattell-Horn-Carroll (CHC) Theory

- Integration of *Gf-Gc* and Three-Stratum theories
- 3 strata, more broad abilities than Three-Stratum theory

Slide 7

Cattell-Horn-Carroll (CHC) Theory



Slide 8

Cattell-Horn-Carroll (CHC) Theory

- Higher-order, mediational structure in which g has indirect effects on measured abilities via second-order abilities
  - Emphasis on lower strata, interpretation of *g* is optional based on theoretical orientation (Schneider & McGrew, 2012)

Slide 9

Cattell-Horn-Carroll (CHC) Theory

- Dominant theory guiding the contemporary, applied assessment of intelligence
    - WJ-IV
    - DAS-2
    - KABC-II
    - SB-5
    - WISC-V

Slide 10

Need for Study

TASP 2016 10

Slide 11

Limitations with Carroll's Analyses

- Relied on exploratory factor analysis (EFA) with Schmid-Leiman (SL) transformations, which did not allow for true bi-factor rotations
  - "SL can only be accurate when certain, highly unlikely, conditions exist (perfect cluster structure, proportionality) and the sample is large enough so that the correlation matrix reflects the population" (Mansolf & Reise, 2016, p. 17)
    - Condition 1: Perfect item structure (items load exclusively on *g* and a single group factor)
    - Condition 2: Proportionality (ratio of general and group factor loadings is the same for all mental tasks associated with a group factor)

TASP 2016 11

Slide 12

Limitations with Carroll's Analyses

- Carroll compared EFA and confirmatory factor analysis (CFA) results for the Gustaffson (1984) and Palmer, Macleod, Hunt, and Davidson (1985) studies
  - Results differed in important ways
  - Carroll argued that the two methods (EFA & CFA) should be used in combination (Carroll, 1995).

TASP 2016 12

Slide 13

Limitations with Carroll's Analyses

- Carroll's placement of abilities into Stratum I or Stratum II was largely a qualitative decision based on re-analysis of 467 studies
  - No single sample has been administered a sufficient range of mental tasks to allow for testing of a model containing all purported abilities
  - Carroll only identified >2 second-order factors in 18 data sets
    - Vast majority (16) of these studies had 3 second-order factors
    - Maximum number of second-order factors identified = 5

Slide 14

Limitations with Carroll's Analyses

- According to Carroll (1993),"Many factors remain inadequately specified, and many aspects of the three-stratum theory need to be tested and refined" (p. 688).

Slide 15

Research Questions

1. Did Carroll over-factor the datasets he analyzed and identify factors that are non-replicable or explain trivial percentages of common factor covariance?
2. To what extent are identified factors sufficiently reliable for clinical interpretation?

Slide 16

Method and Results

Slide 17

Selection of Data Sets

- Focused on 10 studies from which Carroll extracted the most second-order factors
  - Selected to maximize the possibility of identifying Stratum II abilities

Slide 18

Analysis-Study A

- Re-analysis with two methods
  - Jennrich and Bentler’s EFA bi-factor rotation
  - Higher-order EFA with orthogonal transformation
- Comparisons
  - Jennrich and Bentler’s criterion for bi-factor structure, $Q(\boldsymbol{\Lambda})$. Smaller values indicate better bi-factor structure (i.e., loadings on g and 1 other factor).
  - Model-based reliability estimates for each factor
    - Coefficient omega (ω)
    - Omega hierarchical (ωh)

Slide 19

Results-Study A

Study	$\Delta Q(\Lambda)^{\dagger}$	ω		ω_h	
		S-L	BF	S-L	BF
Christal (1958)	-.07	.93	.94	.63	.75
Fogarty (1987)	+.02	.95	.96	.78	.90
Gustafsson (1984)	-.89	.95	.95	.73	.87
Hakstian & Cattell (1974)	-.16	.92	.93	.74	.90
Hakstian & Cattell (1978)	-.16	.81	.87	.47	.82
Horn (1965)	-.59	.94	.94	.67	.90
Horn & Stankov (1982)	-.07	.89	.90	.69	.87
Sung & Dawis (1987)	+.16	.84	.91	.64	.83
Thurstone & Thurstone (1941)	-3.76	.97	.97	.63	.91
Undheim (1981)	-.11	.94	.95	.77	.92

Notes. † = Change in $Q(\Lambda)$ when using an analytic bi-factor rotation rather than the Schmid-Leiman transformation. ω = coefficient omega. ω_h = omega hierarchical. S-L = Schmid-Leiman transformation. BF = bi-factor rotation.

- $Q(\boldsymbol{\Lambda})$ estimates were typically lower when using the bi-factor rotation.
- ω_h was consistently higher for bi-factor models (average for S-L transformation:.68; average for bi-factor rotation: .87).

Slide 20

Results-Study A (cont.)

- The number of well-defined group factors using a bi-factor model typically < higher-order model.
- Group factors more consistent with Stratum I than Stratum II abilities
  - Typically, only two to three tests of similar content had moderate to strong loadings on each group factor.

Slide 21

Analysis-Study B

- Analyzed 5 of 10 previously selected data sets
  - Only data sets for which means and SDs were reported
  - Model for Sung and Dawis (1987) did not converge
- CFA with bi-factor models
  - Initial models based on Carroll's EFA results
  - The Christal (1958) model was bi-factor with correlated unique variances for group factors
    - Correlated unique variances appear to be consistent with Stratum II abilities
    - Akaike weights were used for model comparisons

Slide 22

Results-Study B

- Christal (1958)
  - Identified 9 of 12 factors identified by Carroll (1993)
    - Associative memory, associative memory (color), general information, numerical facility, and motivation (Carroll viewed as Stratum I abilities)
    - Broad visual perception specified as a factor, broad memory ability and crystallized intelligence are represented by correlated group factors (Carroll viewed as Stratum II abilities)
    - *g* (Carroll viewed as Stratum III)

Slide 23

Results-Study B

- Fogarty (1987)
  - Identified 7 of 9 factors identified by Carroll (1993)
    - Spelling ability and time sharing (Carroll viewed as Stratum I abilities)
    - Broad auditory function, broad visual perception, crystallized intelligence, and fluid intelligence (Carroll viewed as Stratum II abilities)
    - *g* (Carroll viewed as Stratum III)

Slide 24

Results-Study B (cont.)

- Hakstian & Cattell(1978)
  - Identified 6 of 8 factors identified by Carroll (1993)
    - Broad memory ability, broad retrieval ability, broad visual perception, crystallized intelligence, and fluid intelligence (Carroll viewed as Stratum II abilities)
    - *g* (Carroll viewed as Stratum III)

Slide 25

Results-Study B (cont.)

- Undheim(1981)
  - Identified 5 of 6 factors identified by Carroll (1993)
    - Broad speediness, broad visual perception, crystallized intelligence, and fluid intelligence (Carroll viewed as Stratum II abilities)
    - *g* (Carroll viewed as Stratum III)

Slide 26

Results-Study B

		g	*Ga*	*Gc*	*Gf*	*Glr*	*Gs*	*Gv*	*Gy*	KO	MA	MA-C	MO	N	SG	TS
	ω	ω_h	ω_s	ω_s	ω_s	ω_s	ω_s	ω_s	ω_s	ω_s	ω_s	ω_s	ω_s	ω_s	ω_s	ω_s
Christal (1958)	.93	.91						.23		.20	.65	.12	.13	.17		
Fogarty (1987)	.96	.96	.41	.13	.13			.19							.19	.27
Hakstian & Cattell (1978)	.85	.83		.06	.17	.17	.10	.07	.05							
Undheim (1981)	.94	.93			.16	.40	.44	.19								

Notes. ***g*** **= general intelligence,** ***Ga*** **= broad auditory function,** ***Gy*** **= broad memory ability,** ***Glr*** **= broad retrieval ability,** ***Gv*** **= broad visual perception,** ***Gc*** **= crystallized intelligence,** ***Gf*** **= fluid intelligence,** ***Gs*** **= broad speediness, KO = general information, MA = associative memory, MO = motivation factor, MA-C = associative memory-color, N = numerical facility, SG = spelling ability, TS = time sharing.**

Slide 27

Results-Study B (cont.)

- Similar to results from Study A
  - The number of well-defined group factors using a bi-factor model typically < higher-order model.
  - Typically, only two to three tests of similar content had moderate to strong loadings on each group factor.
    - Exceptions are associative memory in Christal (1958) and broad auditory function in Fogarty (1987)
  - Estimates of reliability
    - Average for *g* (ω_h) = .91
    - Average for unique variance for Stratum II abilities (ω_s) = .21
    - Average for unique variance for Stratum I abilities (ω_s) = .25

Slide 28

Theoretical Implications

Slide 29

Over-factoring

- Reliance on EFA with SL transformation led to unnecessarily complex theory
  - Some Stratum II and Stratum I abilities likely of little theoretical and/or practical import
  - Most mental tasks examined were found to be good measures of *g*
    - After accounting for *g* there is typically little reliable variance uniquely attributable to group factors

Slide 30

Future Directions

- Results support using bi-factor models rather than higher-order models
  - Guards against over-factoring
- Need for additional investigation regarding the structure of intelligence
- Need for additional investigation to determine what the lower strata abilities explain
  - Theory or taxonomy?
    - Former requires evidence of explaining one or more phenomena

Slide 31

Practical Implications

Slide 32

Carroll's (1993) Goal

- Identify and interpret the abilities that comprise intelligence "without regard" for their relative importance or usefulness (p. 693).

Slide 33

Current Practice

- Interpretation of Stratum II abilities are emphasized in most test manuals
- Interpretations of first- and second-stratum abilities are emphasized in the cross-battery assessment approach (Flanagan, Alfonso, & Ortiz, 2012).
- Results from our analyses do not support citation of Carroll's (1993) work as empirical basis for interpreting lower strata abilities

Slide 34

Acting on Evidence

- With respect to the prediction of educational outcomes, many studies suggest limited incremental validity (beyond *g*) for lower strata abilities
- Absence of evidence of instructional utility for patterns of strengths and weaknesses in lower strata abilities (Miciak et al., 2016)
- Our results provide further evidence against the de-emphasis of *g* in lieu of abilities at lower-strata.

Slide 35

Acting on Evidence

- Brain-behavior isomorphism fallacy (Fletcher & Taylor, 1984)
  - Unclear if performance with behavioral tests reflects neurological dysfunction
  - Cognitive test scores are products of mental activity that reflect individual differences
    - We can make reliable inferences about general ability but not about specific cognitive processes
  - Performance deficits may arise from a variety of sources other than neurological dysfunction

Slide 36

Acting on Evidence

- PSW models
  - Burden for evidenced supporting PSW methods should fall upon those advocating their use (Kranzler et al., 2016)
  - Simulation studies demonstrate limited utility with single indicators of abilities and only modest improvement when using multiple indicators (e.g., Miciak et al., 2014)
  - Difficult to reliably assess strengths and weaknesses due to insufficient unique, reliable variance compounded with imperfect measurement
    - Creates signal to noise problem

Slide 37

Spot the Difference Analogy: Limited variance with low reliability

Slide 38

Spot the Difference Analogy: Limited variance with high reliability

Slide 39

Future Directions

- Test design considerations
  - Possible goal: Maximize variance for *g* (focus on interpretation of *g*)
  - Possible goal: Include tests that maximize unique (non-*g*) variance for group factors
    - Does this unique variance for group factors have utility?
      - Incremental validity for prediction
      - Instructional utility
- Construct scores (Benson et al., 2016)
  - Allows for separation of *g* variance from residual variance for group factors

Slide 40

Questions

Email: Nicholas_Benson@Baylor.edu

Slide 41

References

Benson, N. F., Kranzler, J. H., & Floyd, R. G. (2016). Examining the integrity of measurement of cognitive abilities in the prediction of achievement: Comparisons and contrasts across variables from higher-order and bifactor models. *Journal Of School Psychology*, *58*1-19.

Beaujean, A.A. (2015). John Carroll's Views on Intelligence: Bi-Factor vs. Higher-Order Models. *Journal of Intelligence, 3*, 121-136.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. New York, NY, US: Cambridge University Press. doi:10.1017/CBO9780511571312

Carroll, J. B. (1995). On methodology in the study of cognitive abilities. *Multivariate Behavioral Research*, *30*, 429-452.

Fletcher, J. M., & Taylor, H. G. (1984). Neuropsychological approaches to children: Towards a developmental neuropsychology. *Journal Of Clinical Neuropsychology, 6,* 39-56.

Kranzler, J. H., Floyd, R. G., Benson, N., Zaboski, B., & Thibodaux, L. (2016). Cross-Battery Assessment pattern of strengths and weaknesses approach to the identification of specific learning disorders: Evidence-based practice or pseudoscience?. *International Journal Of School & Educational Psychology*, *4*, 146-157.

Slide 42

References

Kranzler, J. H., Floyd, R. G., Benson, N., Zaboski, B., & Thibodaux, L. (2016). Cross-Battery Assessment pattern of strengths and weaknesses approach to the identification of specific learning disorders: Evidence-based practice or pseudoscience?. *International Journal Of School & Educational Psychology*, *4*, 146-157.

Mansolf, M., & Reise, S. P. (2016). Exploratory bifactor analysis: The Schmid-Leiman orthogonalization and Jennrich-Bentler analytic rotations. Multivariate Behavioral Research, 0, 1-20. http://dx.doi.org/10.1080/00273171.2016.1215898

Miciak, J., Fletcher, J. M., Stuebing, K. K., Vaughn, S., & Tolar, T. D. (2014). Patterns of cognitive strengths and weaknesses: Identification rates, agreement, and validity for learning disabilities identification. *School Psychology Quarterly*, *29*(1), 21-37.

Miciak, J., Williams, J. L., Taylor, W. P., Cirino, P. T., Fletcher, J. M., & Vaughn, S. (2016). Do processing patterns of strengths and weaknesses predict differential treatment response?. *Journal Of Educational Psychology*, *108*, 898-909.

Schneider, W. J., & McGrew, K. S. (2012). The Cattell-Horn-Carroll model of intelligence. In D. P. Flanagan, P. L. Harrison, D. P. Flanagan, P. L. Harrison (Eds.) , *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 99-144). New York, NY, US: Guilford Press.