Research Methods

Confirmatory Factor Analysis Using SPSS Amos

Andy Gillett

1

Factor Analysis

Factor Analysis is a statistical technique

- for analysing the correlations between several variables to reduce them to a smaller number of underlying dimensions, called factors, and
- to determine the correlation of each of the original variables with each factor.

2

Factor Analysis

Factor Analysis takes a large set of variables and looks for a way the data may be 'reduced' or summarised using a smaller set of factors or components.

It does this by looking for clusters or groups among the intercorrelations of a set of variables.

3

Factor Analysis

There are two basic types of Factor Analysis: exploratory and confirmatory.

Exploratory Factor Analysis (EFA) is used when the researcher does not know how many factors are necessary to explain the interrelationships among a set of characteristics, indicators, or items. Therefore, the researcher uses the techniques of Factor Analysis to explore the underlying dimensions of the construct of interest.

4

Confirmatory Factor Analysis

In contrast, Confirmatory Factor Analysis (CFA) is used to assess the extent to which the hypothesized organization of a set of identified factors fits the model. It is used when the researcher has some knowledge about the underlying structure of the construct under investigation.

CFA is used to test hypotheses concerning the structural relationships among a set of factors associated with a specific theory or model.

5

Confirmatory Factor Analysis

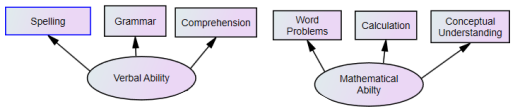
Researchers are assessing the construct validity of a new instrument used to determine aptitude for graduate school. The researchers proposed two latent/not observable factors (Verbal Ability and Mathematical Ability). Each factor has three measured (indicator) variables. The Verbal Ability factor is composed of scores from three subscales: spelling (spell), grammar (gram), and comprehension (comp). The Mathematics Ability factor is composed of scores from three subscales: solving maths word problems (prob), performing calculations (calc), and understanding of mathematical concepts (conc).

(Based on Meyers, Gamst & Guarino, 2017, ch. 11).

6

Confirmatory Factor Analysis

I've used the SPSS data file named VerbalMaths.sav, supplied with IBM SPSS AMOS.

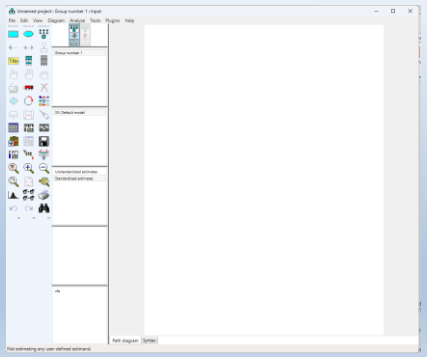


Graduate School Aptitude

7

7

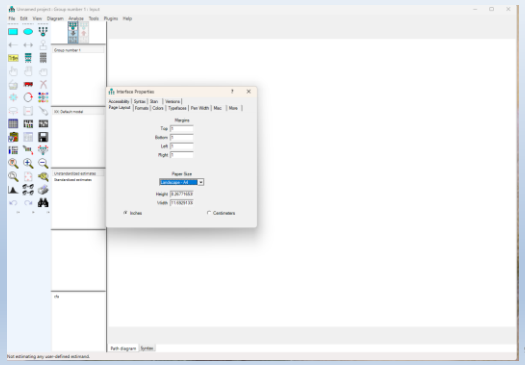
Confirmatory Factor Analysis



8

8

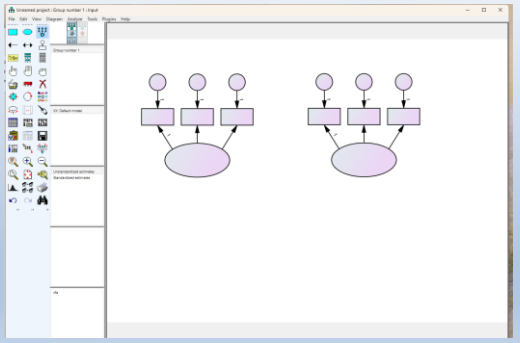
Confirmatory Factor Analysis



9

9

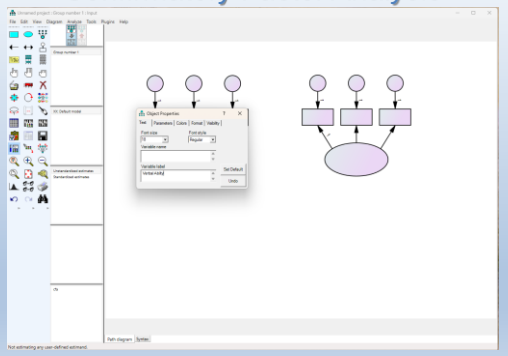
Confirmatory Factor Analysis



10

10

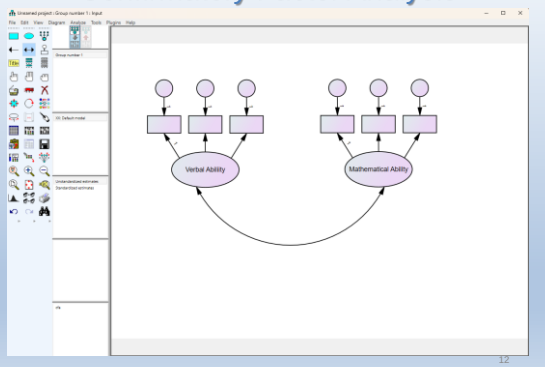
Confirmatory Factor Analysis



11

11

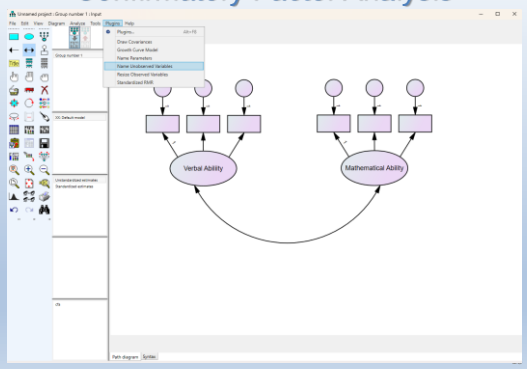
Confirmatory Factor Analysis



12

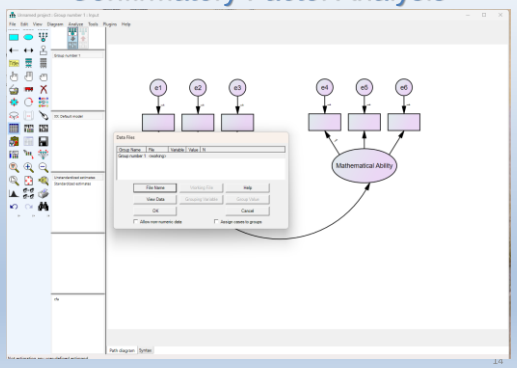
12

Confirmatory Factor Analysis



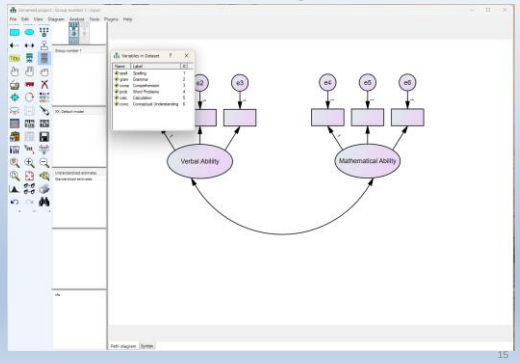
13

Confirmatory Factor Analysis



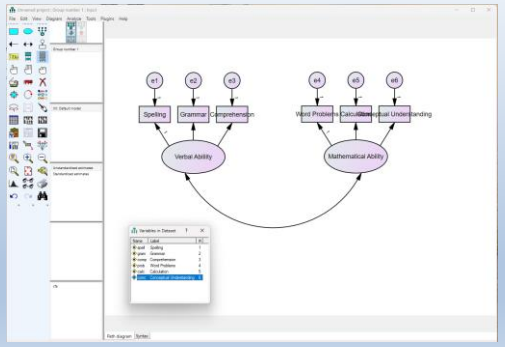
14

Confirmatory



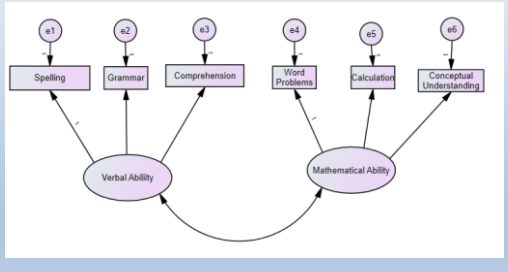
15

Confirmatory Factor Analysis



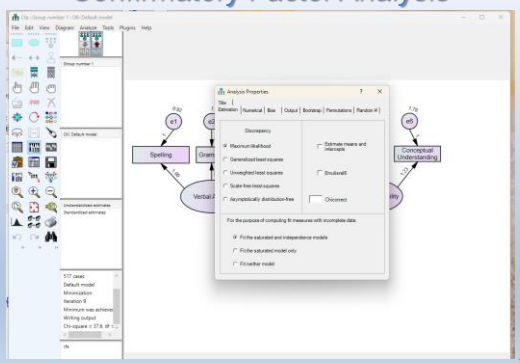
16

Confirmatory Factor Analysis



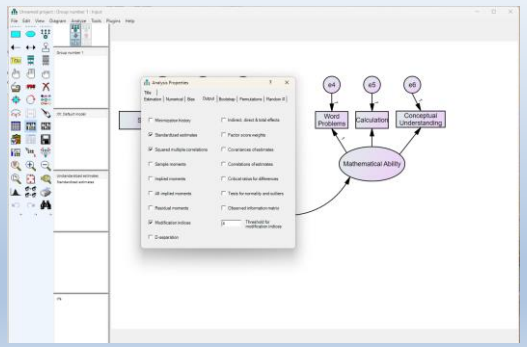
17

Confirmatory Factor Analysis



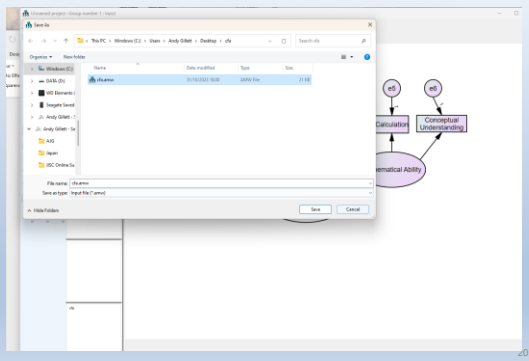
18

Confirmatory Factor Analysis



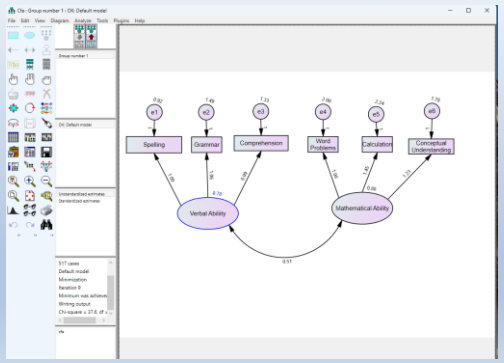
19

Confirmatory Factor Analysis



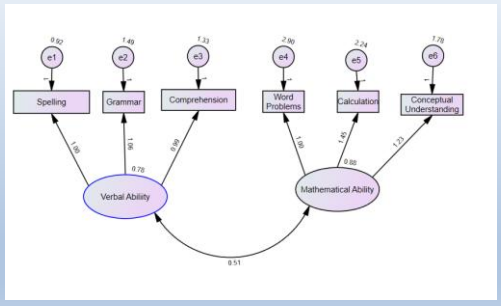
20

Confirmatory Factor Analysis



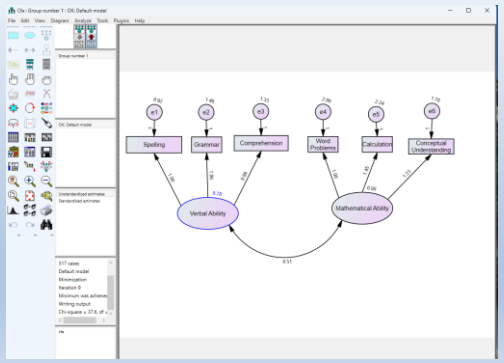
21

Confirmatory Factor Analysis



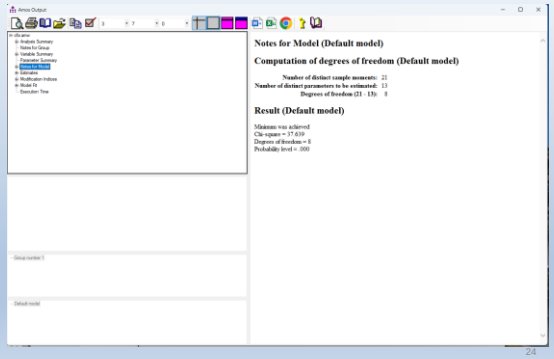
22

Confirmatory Factor Analysis



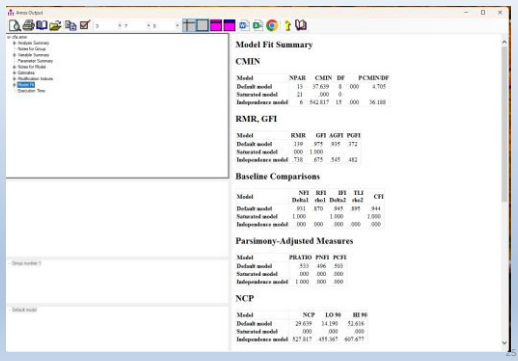
23

Confirmatory Factor Analysis



24

Confirmatory Factor Analysis



Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	13	37.639	8	.000	4.705
Saturated model	21	.000	0		
Independence model	6	542.817	15	.000	36.188

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.139	.975	.935	.372
Saturated model	.000	1.000		
Independence model	.738	.675	.545	.482

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.931	.870	.945	.895	.944
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.533	.496	.503
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	29.639	14.190	52.616
Saturated model	.000	.000	.000
Independence model	527.817	455.367	607.677

25

Confirmatory Factor Analysis

CMIN
CMIN - Chi squared goodness of fit test - want nonsignificant chi squared value - P value <.05 - affected by sample size - for that reason not often used alone
RMR, GFI
GFI - Good Fit Index - equal or greater than 0.90 for a good fit, 0.95 for an excellent fit
Baseline Comparisons
CFI - Comparative Fit Index want between 0.9 (or 0.95) and 1. = good model fit
NFI - Normal Fit Index want between 0.9 (or 0.95) and 1. = good model fit
TLI - Tucker-Lewis Index - want between 0.9 (or 0.95) and 1. = good model fit
RMSEA
RMSEA - Root Mean Square Error of Approximation want 0.05 or below - below 0.08 OK
AIC
AIC - Akaike Information Criterion - the lower the better
BCC - Browne-Cudeck Criterion - the lower the better
ECVI
ECVI = Expected Cross Validation Index - where a smaller value represents a better model fit.

26

Confirmatory Factor Analysis

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	13	37.639	8	.000	4.705
Saturated model	21	.000	0		
Independence model	6	542.817	15	.000	36.188

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.139	.975	.935	.372
Saturated model	.000	1.000		
Independence model	.738	.675	.545	.482

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.931	.870	.945	.895	.944
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

27

Confirmatory Factor Analysis

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.085	.059	.113	.016
Independence model	.261	.243	.280	.000

AIC

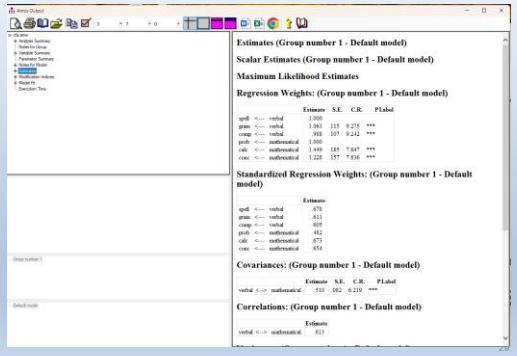
Model	AIC	BCC	BIC	CAIC
Default model	63.639	63.996	118.863	131.863
Saturated model	42.000	42.578	131.209	152.209
Independence model	554.817	554.982	580.305	586.305

ECVI

Model	ECVI	LO 90	HI 90	MECVI
Default model	.123	.093	.168	.124
Saturated model	.081	.081	.081	.083
Independence model	1.075	.935	1.230	1.076

28

Confirmatory Factor Analysis



Estimates (Group number 1 - Default model)

Scalar Estimates (Group number 1 - Default model)

Maximum Likelihood Estimates

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
spell	<---	verbal	1.000				
gram	<---	verbal	1.063	.115	9.275	***	
comp	<---	verbal	.988	.107	9.242	***	
prob	<---	mathematical	1.000				
calc	<---	mathematical	1.449	.185	7.847	***	
conc	<---	mathematical	1.228	.157	7.836	***	

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
spell	<---	verbal	.678
gram	<---	verbal	.611
comp	<---	verbal	.609
prob	<---	mathematical	.682
calc	<---	mathematical	.673
conc	<---	mathematical	.654

Covariances: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
verbal	<-->	mathematical	.510	.082	6.219	***	

Correlations: (Group number 1 - Default model)

			Estimate
verbal	<-->	mathematical	.613

29

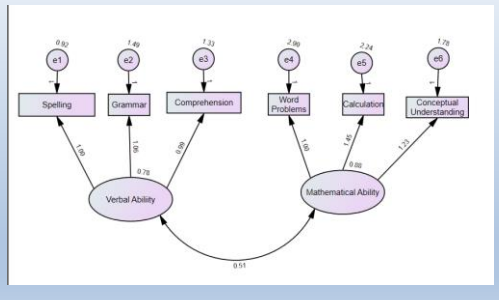
Confirmatory Factor Analysis

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
spell	<---	verbal	1.000				
gram	<---	verbal	1.063	.115	9.275	***	
comp	<---	verbal	.988	.107	9.242	***	
prob	<---	mathematical	1.000				
calc	<---	mathematical	1.449	.185	7.847	***	
conc	<---	mathematical	1.228	.157	7.836	***	

30

Confirmatory Factor Analysis



31

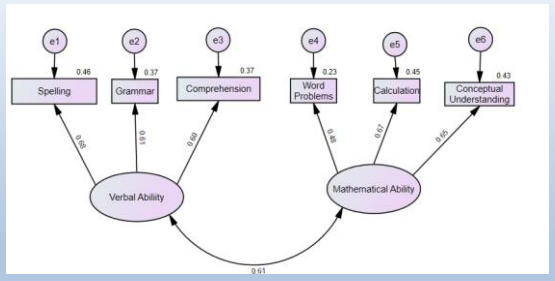
Confirmatory Factor Analysis

Standardized Regression Weights: (Group number 1 - Default model)

			Estimate
spell	<---	verbal	.678
gram	<---	verbal	.611
comp	<---	verbal	.605
prob	<---	mathematical	.482
calc	<---	mathematical	.673
conc	<---	mathematical	.654

32

Confirmatory Factor Analysis



33

Confirmatory Factor Analysis

Covariances: (Group number 1 - Default model)

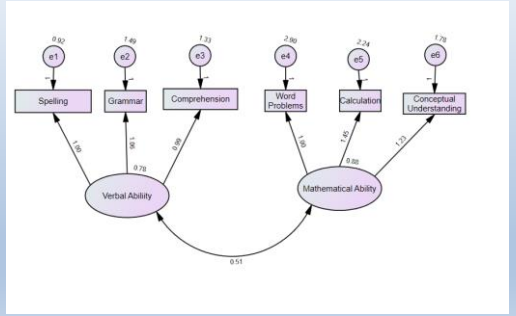
			Estimate	S.E.	C.R.	PLabel
verbal	<-->	mathematical	.510	.082	6.219	***

Correlations: (Group number 1 - Default model)

			Estimate
verbal	<-->	mathematical	.613

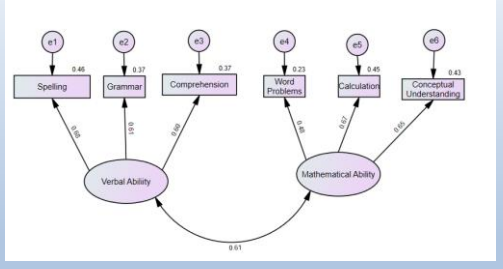
34

Confirmatory Factor Analysis



35

Confirmatory Factor Analysis



36

Confirmatory Factor Analysis

The results of the overall model indicated a moderate fit between the proposed model and the observed data.

The chi squared was statistically significant, indicating a lack of fit.

Both the CFI and NFI indicate a marginal fit with .944 and .931, respectively, because both values are below the recommended .95.

The RMSEA indicates a moderate fit with .085.

The coefficients relating the latent factors to the indicator variables indicate that most of the values are in the .6 range, although the tie between Math Ability and Word Problems is only .48.

37

37

Confirmatory Factor Analysis

The values of these coefficients are well above the generally accepted lower bound of .3 and so would ordinarily be acceptable, but because the fit indexes suggest that the model is less than a good fit to the data, it is useful to look further.

38

38

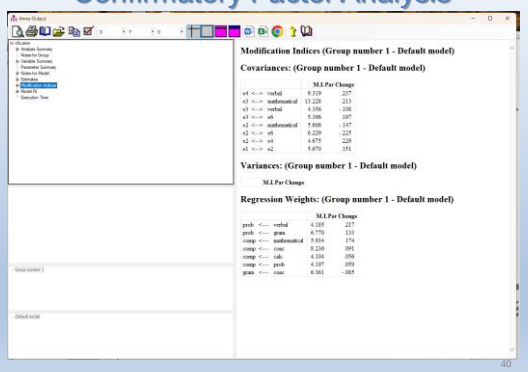
Confirmatory Factor Analysis

Because the fit of the model to the data was not all that impressive, it is possible to speculate that a respecified model (one that is a modified version of the original) may better account for the observed data.

39

39

Confirmatory Factor Analysis



40

40

Confirmatory Factor Analysis

Modification Indices (Group number 1 - Default model)

Covariances: (Group number 1 - Default model)

			M.I.	Par Change
e4	<-->	verbal	9.319	.237
e3	<-->	mathematical	13.228	.213
e3	<-->	verbal	4.356	-.108
e3	<-->	e6	5.396	.197
e2	<-->	mathematical	5.606	-.147
e2	<-->	e6	6.229	-.225
e2	<-->	e4	4.675	.229
e1	<-->	e2	5.670	.151

41

41

Confirmatory Factor Analysis

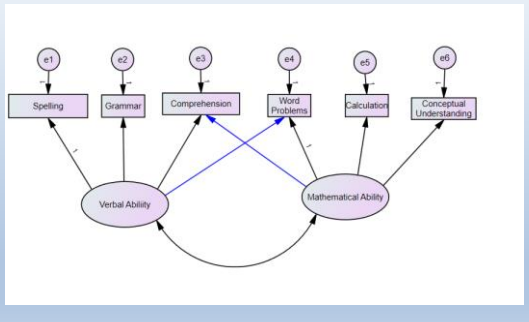
Regression Weights: (Group number 1 - Default model)

			M.I.	Par Change
prob	<---	verbal	4.185	.217
prob	<---	gram	6.770	.133
comp	<---	mathematical	5.934	.174
comp	<---	conc	8.236	.091
comp	<---	calc	4.104	.056
comp	<---	prob	4.197	.059
gram	<---	conc	6.361	-.085

42

42

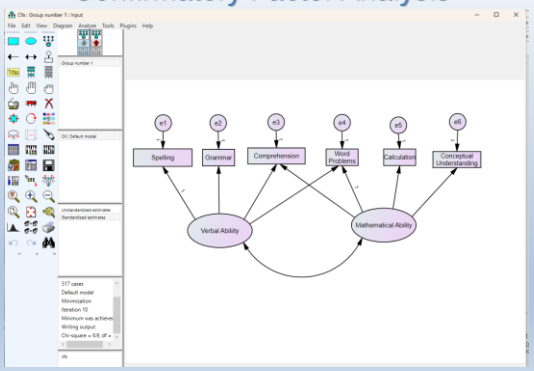
Confirmatory Factor Analysis



43

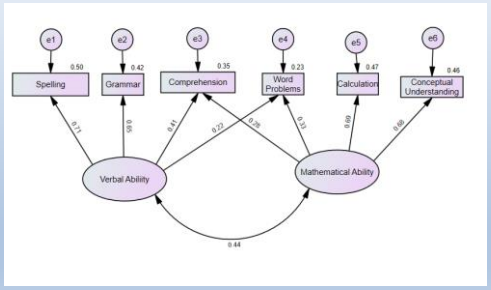
43

Confirmatory Factor Analysis



44

Confirmatory Factor Analysis



45

45

Confirmatory Factor Analysis

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments:	21
Number of distinct parameters to be estimated:	15
Degrees of freedom (21 - 15):	6

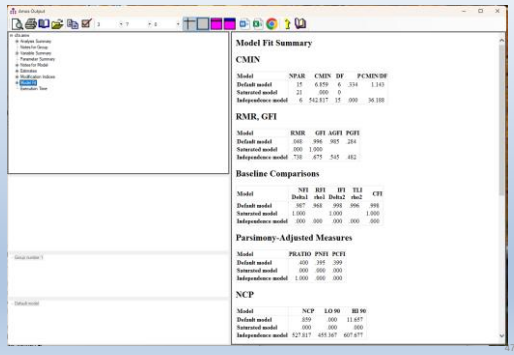
Result (Default model)

Minimum was achieved
Chi-square = 6.859
Degrees of freedom = 6
Probability level = .334

46

46

Confirmatory Factor Analysis



47

47

Confirmatory Factor Analysis

Chi-square and Goodness of Fit Indices for Confirmatory Factor Models

Factor Model	χ^2	df	GFI	NFI	CFI	RMSEA	AIC	BCC	ECVI
Original	37.64*	8	.985	.931	.944	085 (.06 to .11)	63.64	64.00	.12 (.09 to .17)
Respecified	6.859	6	.996	.987	.998	017 (.0 to .061)	36.86	37.27	.07 (.07 to .09)

Note. χ^2 = Chi square test; df= degrees of freedom; GFI = Goodness of Fit Index; NFI = Normed Fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; AIC = Akaike Information Criterion; BCC= Browne-Cudeck Criterion; ECVI = Expected Cross Validation Index.

48

48

Confirmatory Factor Analysis

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
spell	<---	verbal	1.000				
gram	<---	verbal	1.082	.123	8.790	***	
comp	<---	verbal	.638	.106	6.004	***	
prob	<---	mathematical	1.000				
calc	<---	mathematical	2.132	.429	4.973	***	
conc	<---	mathematical	1.842	.370	4.982	***	
comp	<---	mathematical	.625	.165	3.789	***	
prob	<---	verbal	.463	.134	3.446	***	

49

49

Reporting Confirmatory Factor Analysis Results

This study tested the hypothesis that spelling, grammar, and comprehension are indicators of verbal ability and that solving math word problems, performing calculations, and conceptual understanding of mathematical concepts are indicators of math ability. A Confirmatory Factor Analysis of the hypothesized model is shown in Figure 1 and was assessed in IBM SPSS Amos 29. The figure also displays the path coefficients generated in the analysis.

The chi-square and fit indexes are shown in Table 1. Although the chi-square test of the model was statistically significant with a value of 37.64 (8, N = 517), p < .001, the model yielded acceptable fit indexes for the GFI, NFI, and CFI; however, the RMSEA achieved a value of .017 indicating a marginal fit of the model.

Review of the modification indexes led to respecifying the model by including a path from the verbal factor to the measured variable of problem solving (solving math word problems) and including a path from the mathematical factor to the measured variable of comprehension on the grounds that these two skills are tied to each other in the instructional process. Results of the respecified model achieved a nonsignificant chi-square of 6.859.(6, N = 517), p = .334, and higher values on the GFI, NFI, and CFI. Additionally, the RMSEA was lowered to achieve a superior fit (Browne & Cudeck, 1993). The model comparison fit measures, AIC, BCC, and ECVI, were all substantially smaller in the respecified model than the original model, indicating an improved fit between the model and the data.

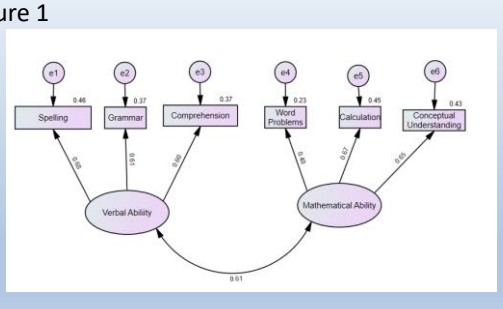
The results suggest that spelling, grammar, and comprehension are indicators of verbal ability but that solving math word problems, performing calculations, conceptual understanding of mathematical concepts, and problem solving are all indicators of math ability. The respecified model is shown in Figure 2.

50

50

Reporting Confirmatory Factor Analysis Results

Figure 1



51

51

Reporting Confirmatory Factor Analysis Results

Table 1

Chi-square and Goodness of Fit Indices for Confirmatory Factor Models

Factor Model	χ^2	df	GFI	NFI	CFI	RMSEA	AIC	BCC	ECVI
Original	37.64*	8	.975	.931	.944	085 (.06 to .11)	63.64	64.00	.12 (.09 to .17)
Respecified	6.859	6	.996	.987	.998	017 (.0 to .061)	36.859	37.272	.071 (.07 to .092)

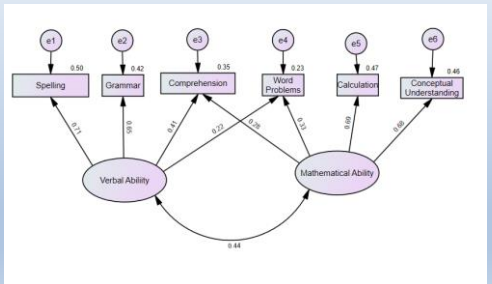
Note. χ^2 = Chi square test; df= degrees of freedom; GFI = Goodness of Fit Index; NFI = Normed Fit Index; CFI = Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; AIC = Akaike Information Criterion; BCC= Browne-Cudeck Criterion; ECVI = Expected Cross Validation Index.

52

52

Reporting Confirmatory Factor Analysis Results

Figure 2



53

53

References & Further Reading

Blunch, N. J. (2008). Introduction to Structural Equation Modelling using SPSS and AMOS. London: Sage.

Burns, R. B. & Burns, R. A. (2008). *Business research methods and statistics using SPSS*. London: Sage.

Byrne, B. M. (2001). Structural equation modeling with AMOS: Basic concepts, applications and programming. Mahwah, NJ: Lawrence Erlbaum.

Catell, R. B. (1966). The scree test for number of factors. *Multivariate Behavioral Research, 1*, 245-76.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* New York: Routledge

Dörnyei, Z. (2007). *Research methods in applied linguistics.* Oxford: Oxford University Press

Gravetter, F. J. & Wallnau, L. B. (1996). Statistics for the behavioral sciences (4th ed). New York: West Publishing Company.

Hand, D. J. (2008) *Statistics: A very short introduction*. Oxford: Oxford University Press.

Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika, 39,* 31-6.

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd ed.). New York: The Guilford Press.

Lowie, W. & Seton, B. (2013). *Essential statistics for applied linguistics*. London: Palgrave Macmillan.

Meyers, L. S., Gamst, G. & Guarino, A. J. (2017). *Applied multivariate research: Design and interpretation* (3rd ed.). London: Sage.

54

54

References & Further Reading

Pallant, J. (2016). *SPSS survival manual* (6th ed.). London: McGraw Hill.

Pett, M. A., Lackey, N. R. & Sullivan, J. J. (2003). *Making sense of Factor Analysis*. London: Sage.

Popham, W. J. , & Sirotnik, K. A. (1973). *Educational statistics: Use and interpretation* (2nd ed.). New York: Harper & Row

Saunders, M. & Lewis, P. (2012). *Doing research in business and management*. London: FT/Prentice Hall.

Saunders, M., Lewis, P. & Thornhill, A. (2016). *Research methods for business students* (7th ed.). London: FT/Prentice Hall.

Shavelson, R. (1988). *Statistical reasoning for the behavioral sciences*. Boston: Allyn and Bacon.

Spiegelhalter, D. (2019). *The art of statistics: Learning from data.* London: Pelican.

Tabachnick, B. G. & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston: Allyn & Bacon.

Thurstone, L. L. (1947). *Multiple Factor Analysis*. Chicago: University of Chicago Press.

Woodrow, L. J. (2006). A model of adaptive language learning. *Modern Language Journal, 90*, 297-319.

Woodrow, L. J. (2010). Researching motivation. In B. Paltridge & A. Phakiti (Eds.), *Continuum companion to research methods in applied linguistics* (pp. 301–317). London: Continuum.

55