8th JACET Vocabulary Acquisition Research Group Conference (December 10, 2011)

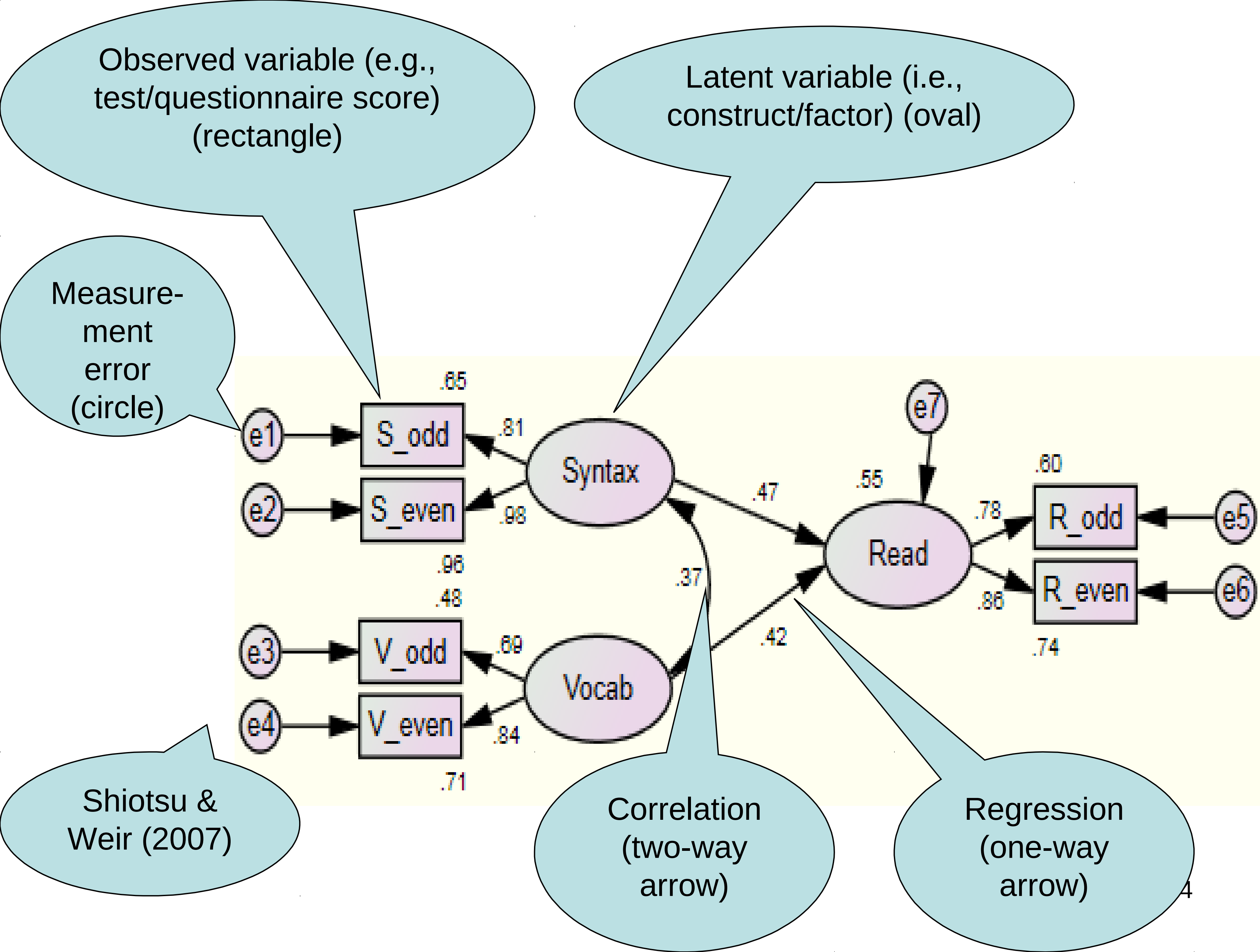
An introduction to structural equation modeling for vocabulary research

Yo In’nami
Toyohashi University of Technology
innami@las.tut.ac.jp
www7b.biglobe.ne.jp/~koizumi/Innami/top-english.html

Overview

- **SEM basics**
- SEM demo
- Applications

- Structural equation modeling (SEM)
  - Also called covariance structure analysis or simultaneous equation modeling
  - A statistical technique for examining the nature of the relationships among observed and latent variables that applies a confirmatory, hypothesis-testing approach to the data (e.g., Byrne, 2006)
  - Regression + factor analysis
  - Encompasses ANOVA, ANCOVA, CFA, regression…
  - Suitable for visually presenting study findings

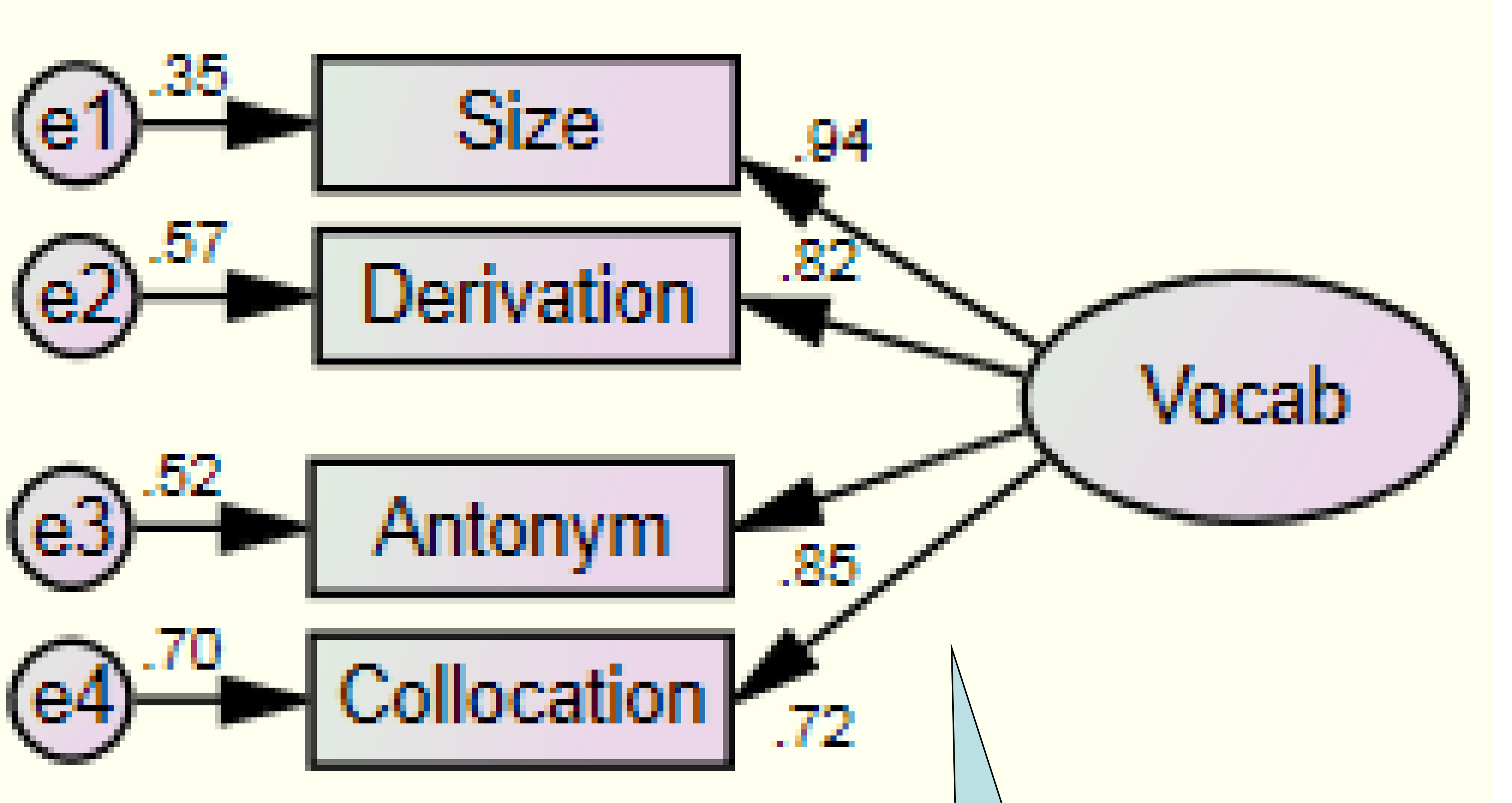

Observed variable (e.g., test/questionnaire score) (rectangle)
Latent variable (i.e., construct/factor) (oval)
Measure-ment error (circle)
Shiotsu & Weir (2007)
Correlation (two-way arrow)
Regression (one-way arrow)
e1
e2
e3
e4
e5
e6
e7
S_odd
S_even
V_odd
V_even
R_odd
R_even
Syntax
Vocab
Read
.65
.81
.98
.96
.48
.69
.84
.71
.37
.47
.42
.55
.60
.78
.86
.74

The diagram was displayed at the workshop.

Tseng, Dörnyel, & Schmitt (2006, p. 93)

The diagram was displayed at the workshop.

Mizumoto & Takeuchi (in press)



Koizumi & In'nami (in preparation)

- Four advantages of SEM (Byrne, 2006)
  - SEM takes a **confirmatory, hypothesis-testing approach** to the data, in contrast to traditional analysis, such as exploratory factor analysis, where analysis is data driven.
  - SEM is designed to **correct for measurement errors** of variables. The results allow a researcher to interpret the relationship among variables, separating the measurement errors.
  - SEM can **analyze both unobserved (i.e., latent) and observed variables**. This contrasts with path analysis that enables researchers to model only observed variables. Latent variables are used to define factors or constructs.
  - **Multivariate relations or indirect effects can be analyzed** using SEM, whereas no other statistical methods can easily do this. Investigation into multivariate relations may include models where correlations are hypothesized only among a certain set of variables. Investigating indirect effects may include determining whether an independent variable directly affects a dependent variable or whether it does so through a mediating variable. Path analysis can be used to model these multivariate relations or indirect effects with observed variables, but it cannot be used to conduct analyses using unobserved variables.

- Five steps involved in an SEM application (Bollen & Long, 1993)
  - Model specification
  - Model identification
  - Parameter estimation
    - E.g., Maximum likelihood
  - Model fit
    - E.g., CFI
  - Model respecification

- Requirements:
  - Sample size: 100–200+
  - Normality
    - Univariate skewness & kurtosis
    - Multivariate kurtosis
  - Missing data

Overview

- SEM basics
- **SEM demo**
- Applications

SEM demo

- Koizumi & In'nami (in preparation)
  - Examining the uni-factor structure of vocabulary knowledge
  - SPSS & Amos
  - Sample size: 100–200+ (224)
  - Normality
    - Univariate skewness & kurtosis (OK)
    - Multivariate kurtosis (OK)
  - Missing data (No missing data)

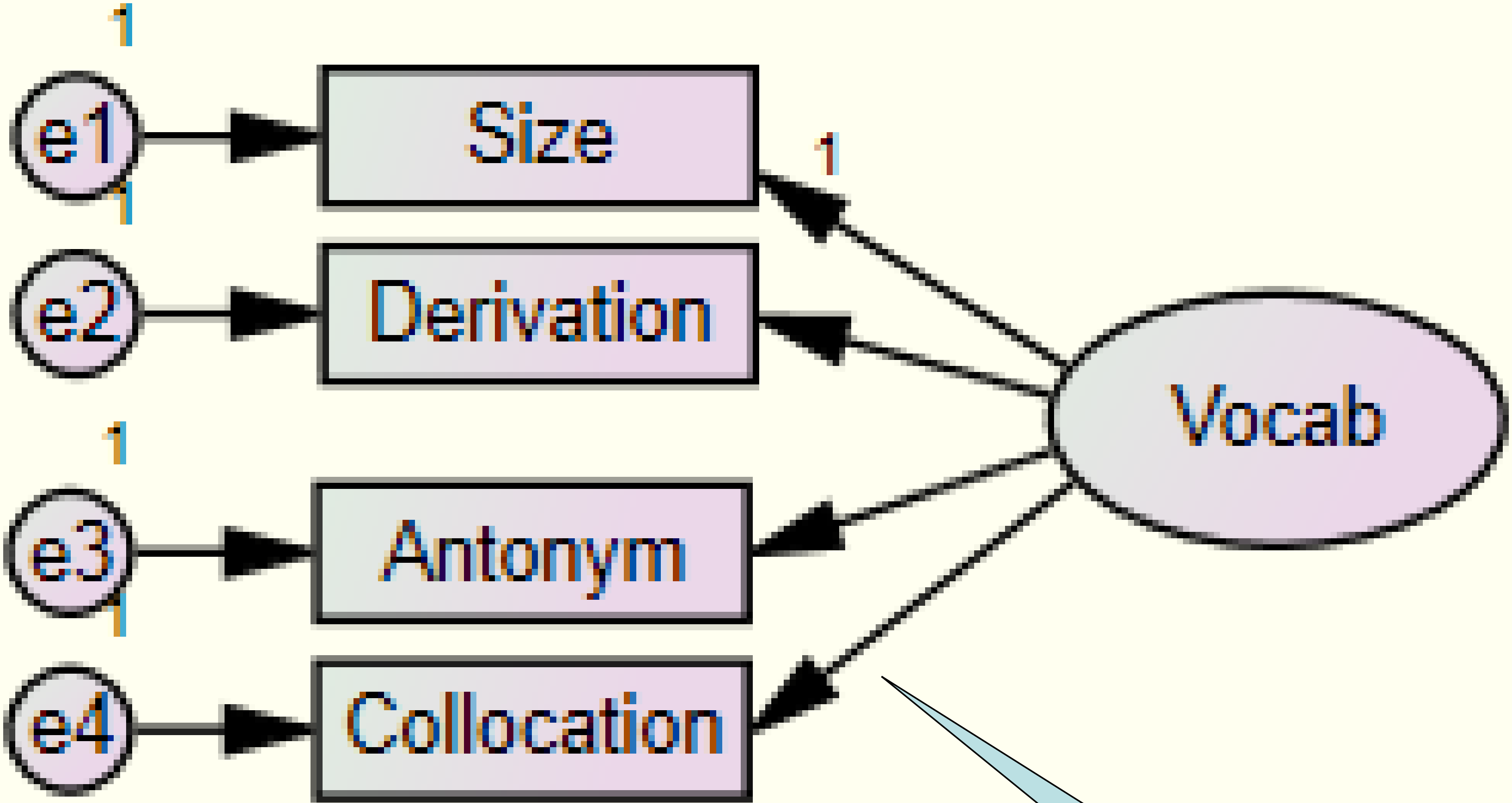
111119Yo_voc_workshop.sav [データセット1] - PASW Statistics データ エディタ

ファイル(F) 編集(E) 表示(V) データ(D) 変換(T) 分析(A) グラフ(G) ユーティリティ(U) ウィンドウ(W)

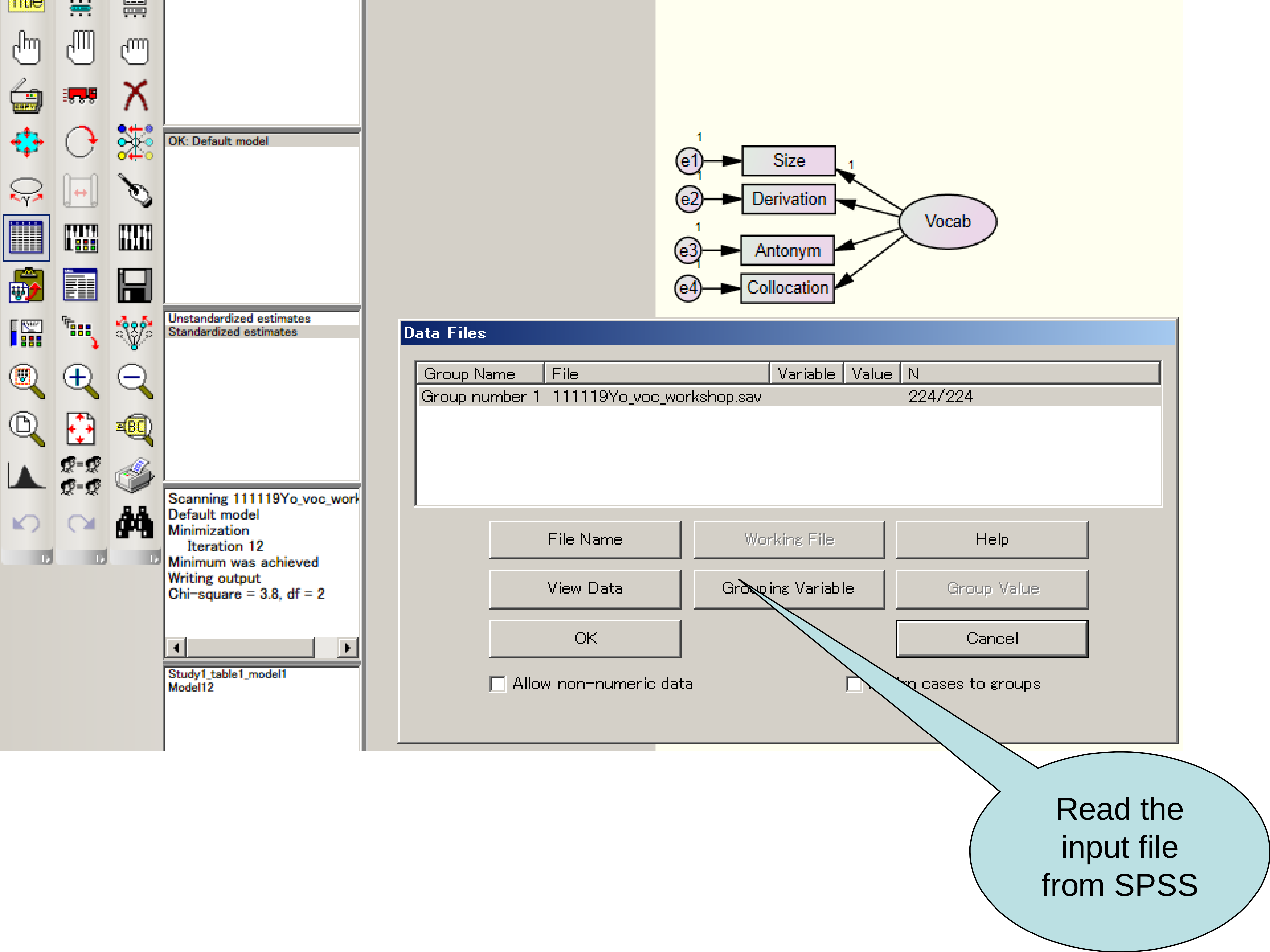
1 : size1 12.00

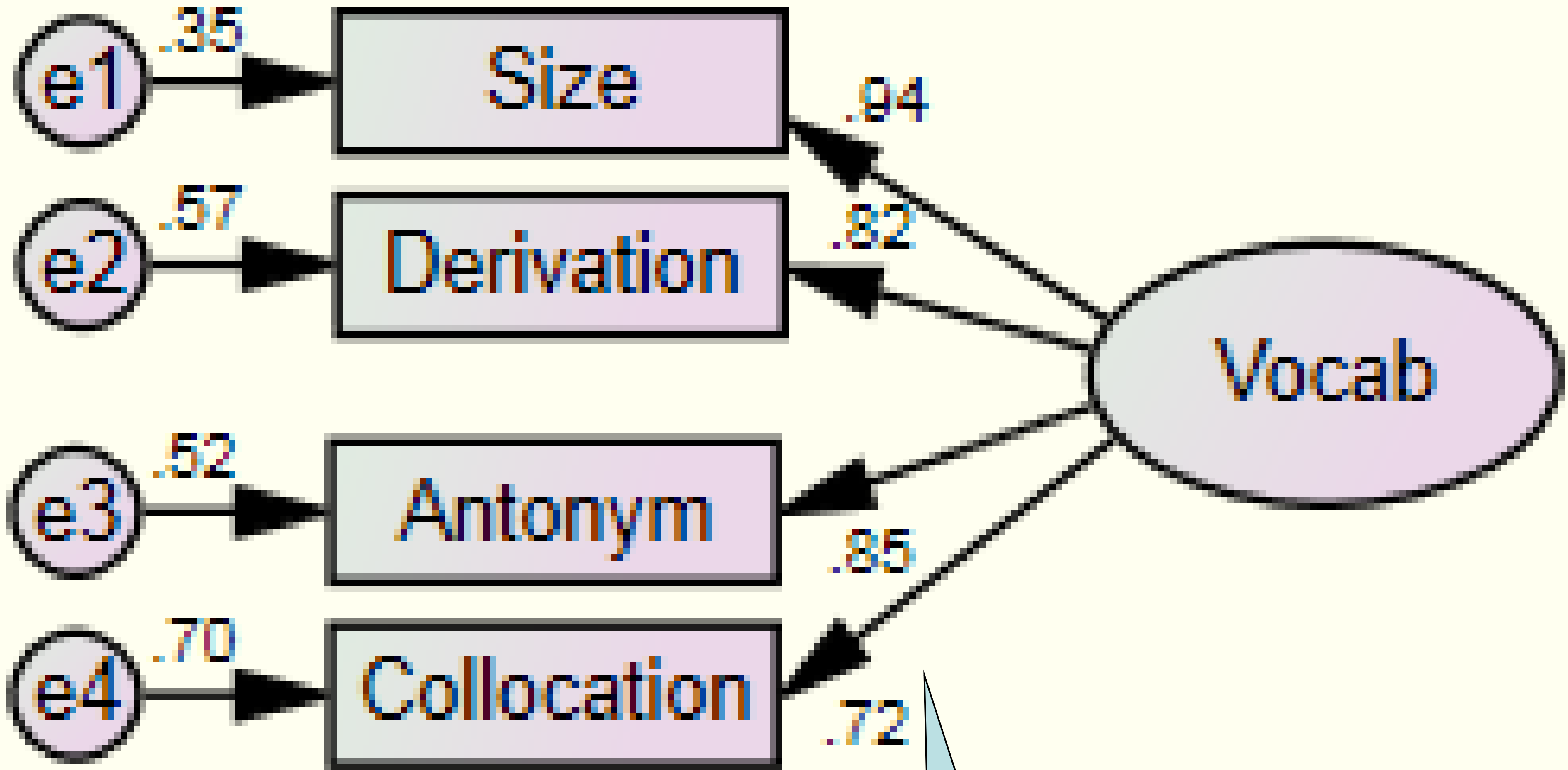
	size1	size2	size3	size	derivation	antonym	collocation
1	12.00	7.00	2.00	21.00	4.00	5.00	8.00
2	7.00	.0	1.00	8.00	3.00	2.00	8.00
3	10.00	5.00	1.00	16.00	4.00	4.00	4.00
4	8.00	3.00	3.00	14.00	6.00	2.00	10.00
5	8.00	2.00	1.00	11.00	2.00	4.00	6.00
6	21.00	16.00	3.00	40.00	10.00	8.00	11.00
7	19.00	9.00	2.00	30.00	8.00	10.00	7.00
8	23.00	18.00	11.00	52.00	13.00	13.00	11.00
9	8.00	2.00	1.00	11.00	2.00	2.00	6.00
10	21.00	11.00	4.00	36.00	9.00	7.00	13.00
11	13.00	9.00	2.00	24.00	6.00	9.00	11.00
12	20.00	15.00	8.00	43.00	10.00	8.00	14.00
13	22.00	15.00	7.00	44.00	10.00	9.00	12.00
14	15.00	7.00	2.00	24.00	6.00	7.00	6.00
15	6.00	5.00	1.00	12.00	2.00	3.00	5.00
16	19.00	10.00	3.00	32.00	12.00	6.00	11.00
17	18.00	10.00	2.00	30.00	10.00	8.00	[illegible]
18	18.00	11.00	2.00	31.00	8.00	8.00	13.00
19	16.00	13.00	2.00	31.00	9.00	8.00	7.00
20	10.00	5.00	2.00	17.00	3.00	4.00	11.00

Create the input data file in SPSS/Excel



Draw the model using Amos


OK: Default model
Unstandardized estimates
Standardized estimates
Scanning 111119Yo_voc_work
Default model
Minimization
Iteration 12
Minimum was achieved
Writing output
Chi-square = 3.8, df = 2
Study1_table1_model1
Model12
e1
e2
e3
e4
Size
Derivation
Antonym
Collocation
Vocab
Data Files
Group Name
File
Variable
Value
N
Group number 1
111119Yo_voc_workshop.sav
224/224
File Name
Working File
Help
View Data
Grouping Variable
Group Value
OK
Cancel
Allow non-numeric data
cases to groups
Read the input file from SPSS



Run the model

Assessment of normality (Group number 1)

Variable	min	max	skew	c.r.	kurtosis	c.r.
antonym	.000	14.000	.143	.871	-.486	-1.484
collocation	2.000	18.000	-.353	-2.159	.533	1.627
size	8.000	54.000	.219	1.337	-.535	-1.636
derivation	.000	19.000	.066	.401	-.531	-1.622
Multivariate					-.372	-.402

Univariate normality: (1) Skewness & kurtosisis, c.r. (critical ratio) ≤ ±1.96 (or 3.29). (2) Better to examine the histogram and the skewness & kurtosis statistics rather than to calculate their significance (N ≥ 200; Field, 2005, p. 72).

Multivariate normality: (1) c.r. (critical ratio) ≤ ±1.96 (or 3.29), (2) c.r. values > 5.00 indicate nonnormal distribution (Bentler, 2005, p. 106; Byrne, 2010, p. 104).

Observations farthest from the centroid (Mahalanobis distance) (Group number 1)

Observation number	Mahalanobis d-squared	p1	p2
30	13.715	.008	.844
201	13.313	.010	.648
57	12.381	.015	.642
168	12.123	.016	.504
158	11.760	.019	.431
209	11.252	.024	.445
[illegible]	11.057	.026	.363
173	10.958	.027	.262
160	9.728	.045	.688
133	9.603	.048	.628
34	9.599	.048	.506
172	9.554	.049	.408
198	9.486	.050	.332

Multivariate normality: Mahalanobis distance less than 13.816 (for *df* = 2, $p < .001$, χ^2 = 13.816)

Models

Default model (Default model)

Notes for Model (Default model)

Computation of degrees of freedom (Default model)

Number of distinct sample moments:	10
Number of distinct parameters to be estimated:	8
Degrees of freedom (10 - 8):	2

1 or above, good; if negative, the model can’t be tested.

Result (Default model)

Minimum was achieved

Chi-square = 3.846

Degrees of freedom = 2

Probability level = .146

Regression Weights: (Group number 1 - Default model)

			Estimate	S.E.	C.R.	P	Label
derivation	<---	e2	2.074	.121	17.147	***	
size	<---	e1	3.595	.408	8.812	***	
collocation	<---	Vocab	.214	.016	13.337	***	
collocation	<---	e4	1.974	.103	19.251	***	
antonym	<---	Vocab	.261	.014	18.217	***	
antonym	<---	e3	1.521	.096	15.890	***	
size	<---	Vocab	1.000				
derivation	<---	Vocab	.316	.019	17.047	***	

All paths are statistically significant.

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

			Estimate
derivation	<---	e2	.567
size	<---	e1	.353
collocation	<---	Vocab	.719
collocation	<---	e4	.695
antonym	<---	Vocab	.853
antonym	<---	e3	.522
size	<---	Vocab	.936
derivation	<---	Vocab	.824

Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	8	3.846	2	.146	1.923
Saturated model	10	.000	0		
Independence model	4	586.794	6	.000	97.799

RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.193	.992	.958	.198
Saturated model	.000	1.000		
Independence model	13.709	.407	.011	.244

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.993	.980	.997	.990	.997
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Parsimony-Adjusted Measures

Model	PRATIO	PNFI	PCFI
Default model	.333	.331	.332
Saturated model	.000	.000	.000
Independence model	1.000	.000	.000

NCP

Model	NCP	LO 90	HI 90
Default model	1.846	.000	11.625
Saturated model	.000	.000	.000
Independence model	580.794	504.966	664.020

FMIN

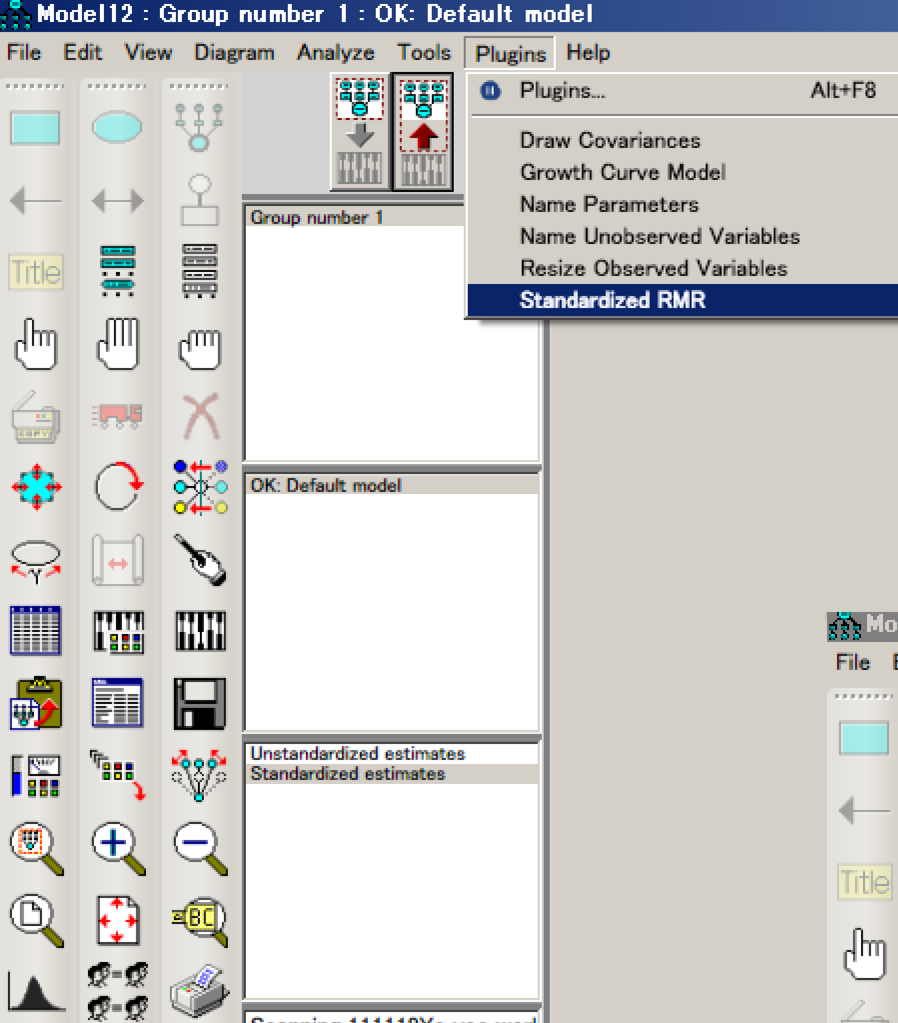
Model	FMIN	F0	LO 90	HI 90
Default model	.017	.008	.000	.052
Saturated model	.000	.000	.000	.000
Independence model	2.631	2.604	2.264	2.978

RMSEA

Model	RMSEA	LO 90	HI 90	PCLOSE
Default model	.064	.000	.161	.299
Independence model	.659	.614	.704	.000

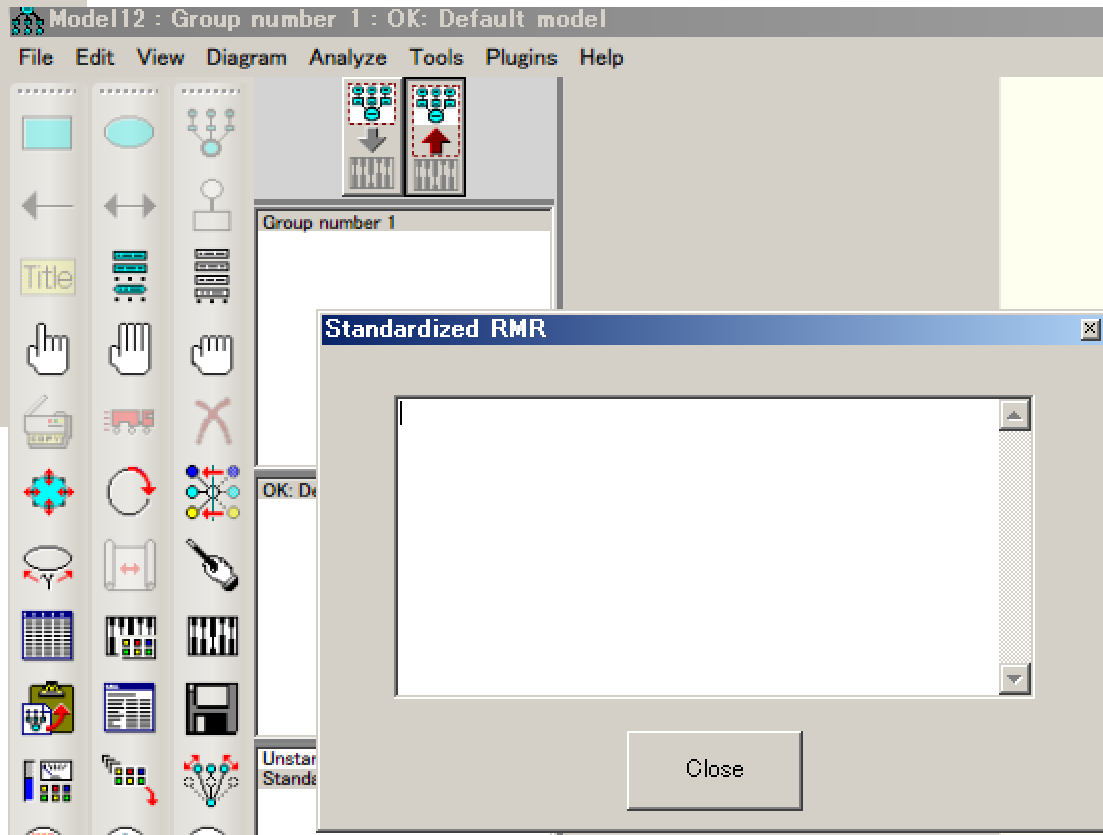
	X^2 (CMIN)	*df*	*p*	CFI	TLI	RMSEA (90% CI)	$p_{\text{close-fit H0}}$	SRMR
Our Model	3.846	2	.146	.997	0.990	0.064 (0.000, 0.161)	.299	.014
Criteria			nonsig	> .95	Near 1.00	=< 0.05	> .05	=< .08

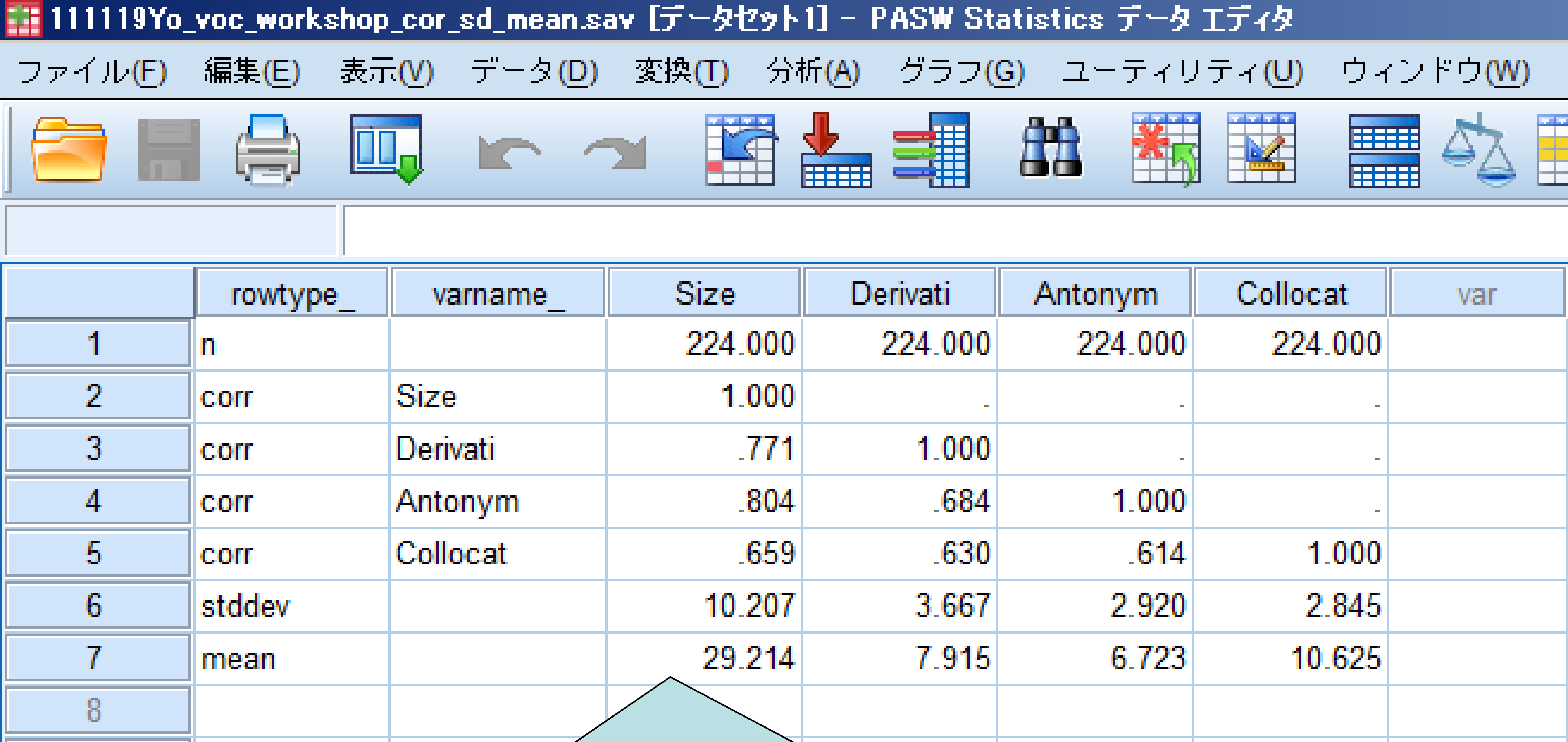
This tests the null hypothesis that the population RMSEA is no greater than .05.



How to calculate SRMR

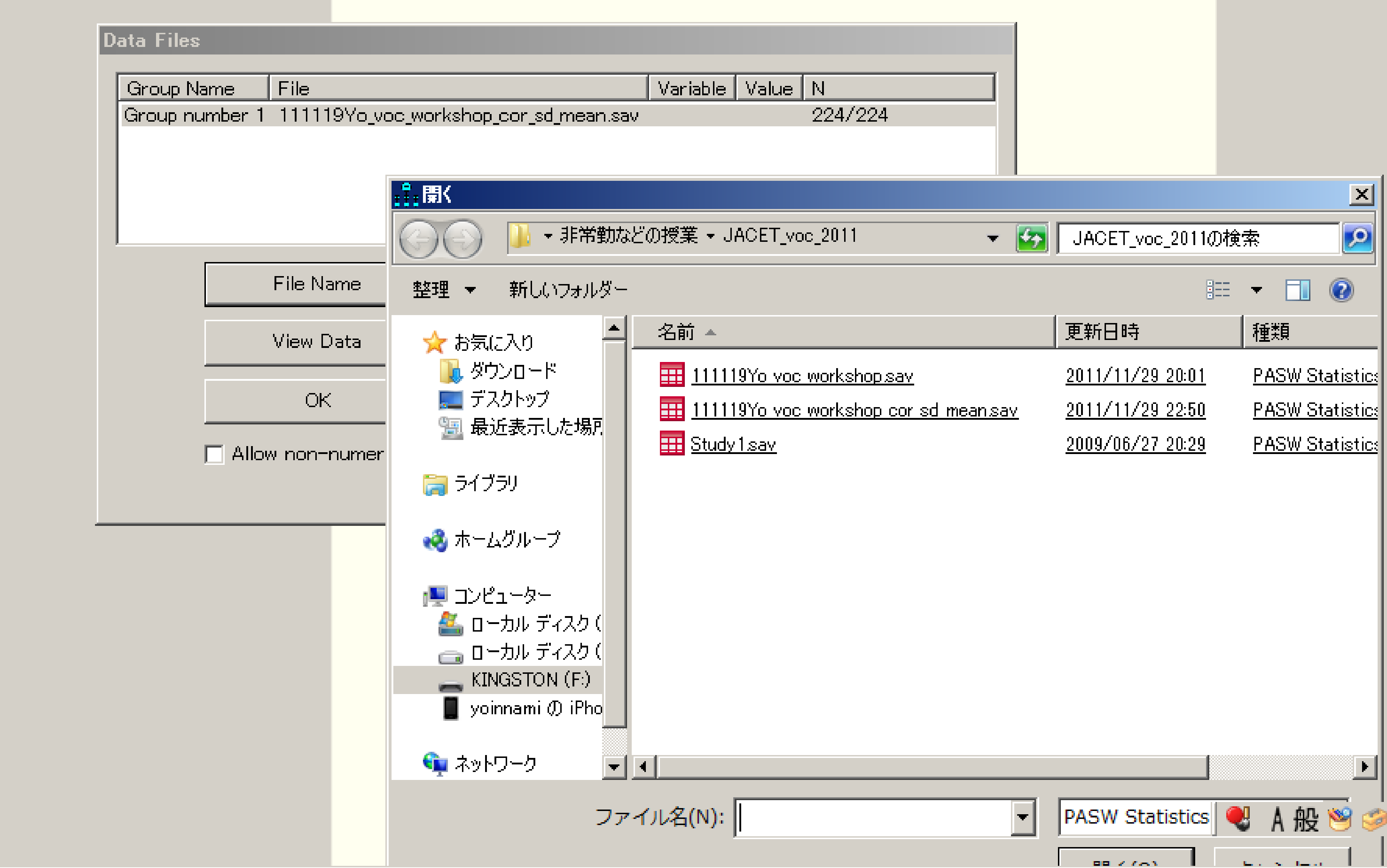
- Run the model with the SRMR box left open and blank.

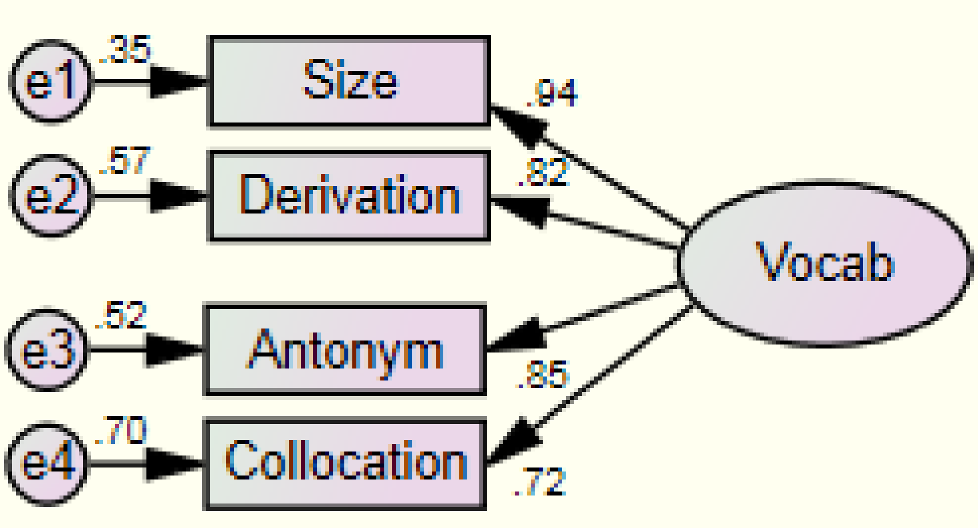




	rowtype_	varname_	Size	Derivati	Antonym	Collocat	var
1	n		224.000	224.000	224.000	224.000	
2	corr	Size	1.000	.	.	.	
3	corr	Derivati	.771	1.000	.	.	
4	corr	Antonym	.804	.684	1.000	.	
5	corr	Collocat	.659	.630	.614	1.000	
6	stddev		10.207	3.667	2.920	2.845	
7	mean		29.214	7.915	6.723	10.625	
8							

SEM results are (generally) reproducible even without the raw data, given access to (1) correlations & *SD*s [+means]) or (2) variances/covariances. This suggests that we can <u>reproduce previous studies</u> to see if the model was correctly analyzed and/or examine alternative models not tested in the primary study (see In'nami & Koizumi, 2010, for further details).


Data Files
Group Name
File
Variable
Value
N
Group number 1
111119Yo_voc_workshop_cor_sd_mean.sav
224/224
File Name
View Data
OK
Allow non-numer
開く
非常勤などの授業
JACET_voc_2011
JACET_voc_2011の検索
整理
新しいフォルダー
お気に入り
ダウンロード
デスクトップ
最近表示した場所
ライブラリ
ホームグループ
コンピューター
ローカル ディスク (
ローカル ディスク (
KINGSTON (F:)
yoinnami の iPho
ネットワーク
名前
更新日時
種類
111119Yo voc workshop.sav
2011/11/29 20:01
PASW Statistics
111119Yo voc workshop cor sd mean.sav
2011/11/29 22:50
PASW Statistics
Study1.sav
2009/06/27 20:29
PASW Statistics
ファイル名(N):
PASW Statistics



CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	8	3.846	2	.146	1.923
Saturated model	10	.000	0		
Independence model	4	586.794	6	.000	97.799

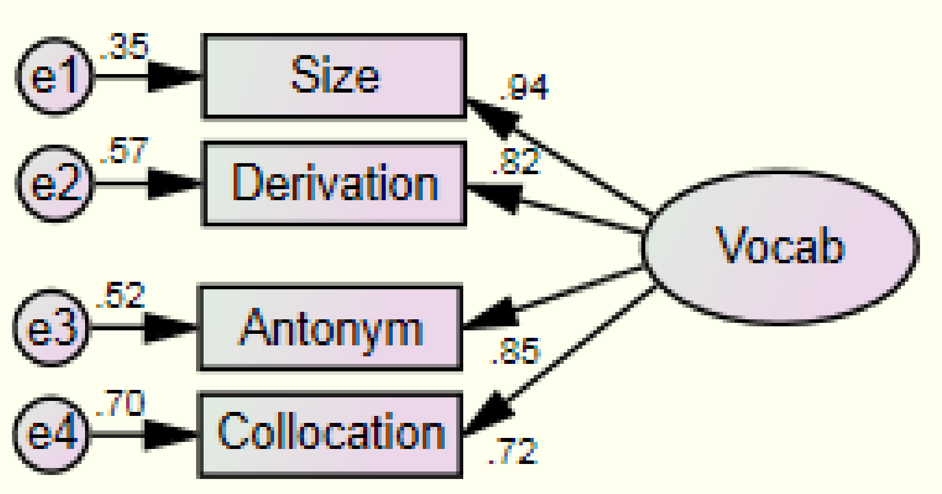
RMR, GFI

Model	RMR	GFI	AGFI	PGFI
Default model	.193	.992	.958	.198
Saturated model	.000	1.000		
Independence model	13.709	.407	.011	.244

Baseline Comparisons

Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.993	.980	.997	.990	.997
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Left; raw data input. Right; *r*, *M*, *SD* input



Model Fit Summary

CMIN

Model	NPAR	CMIN	DF	P	CMIN/DF
Default model	8	3.902	2	.142	1.951
Saturated model	10	.000	0		
Independence model	4	586.777	6	.000	97.796

RMR, GFI

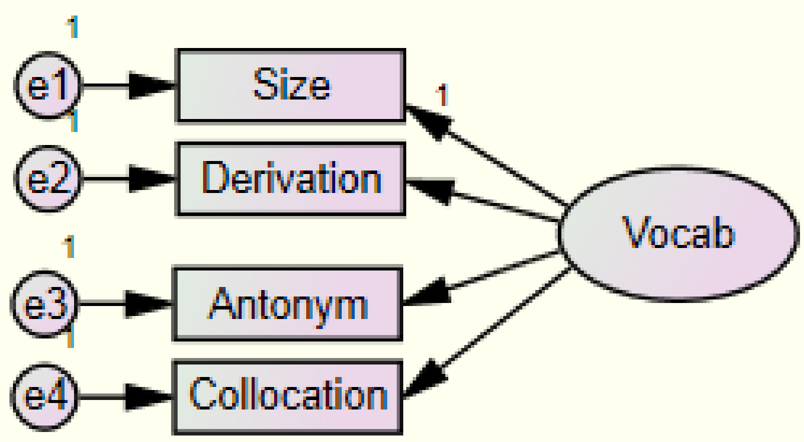
Model	RMR	GFI	AGFI	PGFI
Default model	.194	.991	.957	.198
Saturated model	.000	1.000		
Independence model	13.708	.407	.011	.244

Baseline Comparisons

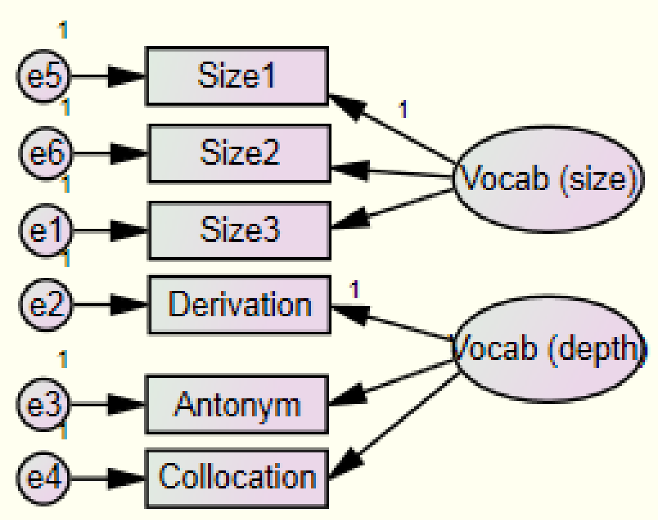
Model	NFI Delta1	RFI rho1	IFI Delta2	TLI rho2	CFI
Default model	.993	.980	.997	.990	.997
Saturated model	1.000		1.000		1.000
Independence model	.000	.000	.000	.000	.000

Overview

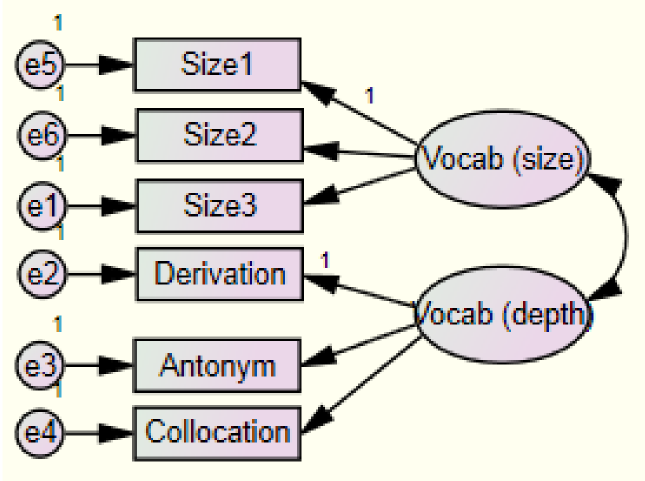
- SEM basics
- SEM demo
- **Applications**



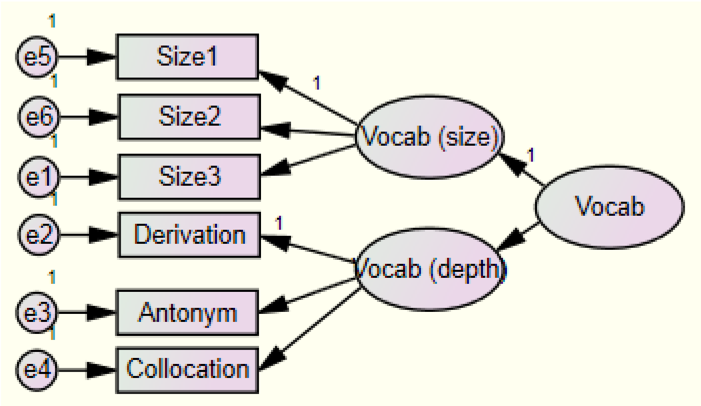
Cross-sectional investigation into growth

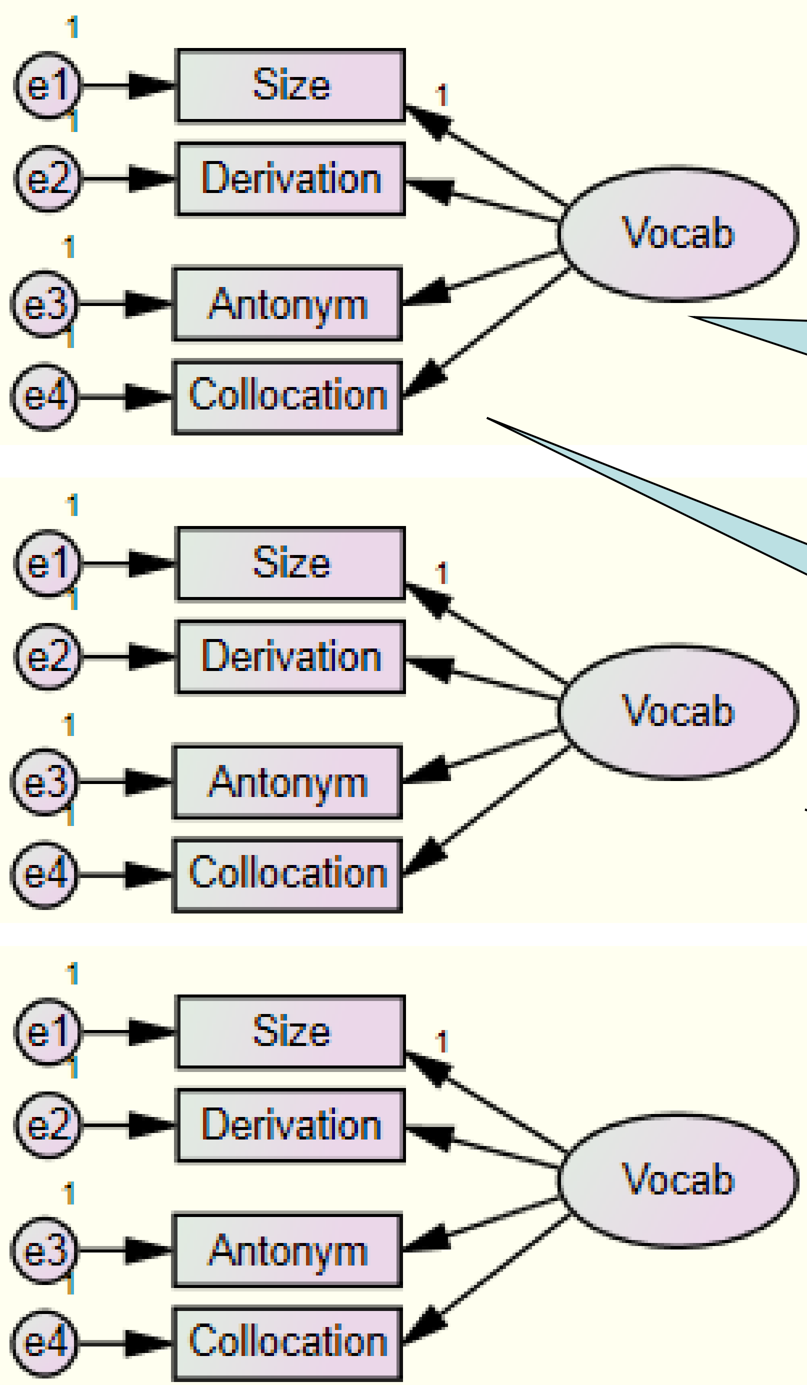


Is there a high correlation between size and depth (Is vocabulary knowledge a unitary construct)? Is the structure different across learners of different proficiency (novice, intermediate, advanced)? (multi-sample model)



Does the strength of association between size and depth vary such that a model of separate size and depth is more appropriate? (hierarchical/higher-order model)





Longitudinal investigation into growth

Does this model hold as learners expand their vocabulary knowledge? Are there any moderating variables? (latent growth model)

Year 1

Year 2

Year 3

- The diagram was displayed at the workshop.

Tseng & Schmitt, (2008)

- The diagram was displayed at the workshop.

Tremblay & Gardner (1995)

- The diagram was displayed at the workshop.

Yashima (2002)

- The diagram was displayed at the workshop.

Matsumura (2003);
latent growth model

- The diagram was displayed at the workshop.

Llosa (2007);
hierarchical/higher-order model

- The tables were displayed at the workshop.

In’nami & Koizumi, 2011

- The table was displayed at the workshop.

Software

	GUI	Detailed manual	Fast & dependable technical support	Missing data	Nonnormal data
Amos	x			x	(x)
EQS	x	xx	xxx	x	x
LISREL	x		x	x	x
Mplus		x	xxx	x	x

- GUI = graphic user interface.

SEMNET

- http://www2.gsu.edu/~mkteer/semnet.html

References

- Bentler, P. M. (2005). *EQS 6 structural equations program manual.* Encino, CA: Multivariate Software.
- Bollen, K. A., & Long, J. S. (1993). Introduction. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 1–9). Newbury Park, CA: Sage.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research.* New York: Guilford.
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Mahwah, NJ: Erlbaum.
- Byrne, B. M. (2006). *Structural equation modeling with EQS: Basic concepts, applications, and programming* (2nd ed.). Mahwah, NJ: Erlbaum.
- Byrne, B. M. (2011). *Structural equation modeling with Mplus: Basic concepts, applications, and programming.* Mahwah, NJ: Erlbaum.
- Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.
- In'nami, Y., & Koizumi, R. (2010). Can structural equation models in second language testing and learning research be successfully replicated? *International Journal of Testing, 10,* 262–273.
- In'nami, Y., & Koizumi, R. (2011). Structural equation modeling in language testing and learning research: A review. *Language Assessment Quarterly, 8,* 250–276.

References

- Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). New York: Guilford.
- Koizumi, R., & In'nami, Y. (in preparation).
- Llosa, L. (2007). Validating a standards-based classroom assessment of English proficiency: A multitrait-multimethod approach. *Language Testing, 24,* 489–515.
- Matsumura, S. (2003). Modelling the relationships among interlanguage pragmatic development, L2 proficiency, and exposure to L2. *Applied Linguistics, 24,* 465–491.
- Mizumoto, A., & Takeuchi, O. (in press). Adaptation and validation of self-regulating capacity in vocabulary learning scale. *Applied Linguistics.*
- Shiotsu, T., & Weir, C. J. (2007). The relative significance of syntactic knowledge and vocabulary breadth in the prediction of reading comprehension test performance. *Language Testing, 24,* 99–128.
- Tremblay, P. F., & Gardner, R. C. (1995). Expanding the motivation construct in language learning. *Modern Language Journal, 79,* 505–518.
- Tseng, W.-T., Dornyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics, 27,* 78–102.
- Tseng, W.-T., & Schmitt, N. (2008). Toward a model of motivated vocabulary learning: A structural equation modeling approach. *Language Learning*, *58*, 357–400.
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *Modern Language Journal, 86,* 54–66.