Assessing the Dimensionality of the Stanford English Language Proficiency Test (SELP):

Factorial Structure and Invariance across two Different Groups of Immigrant Students in the U.S.

Presented by Anja Römhild
Buros Center for Testing - University of Nebraska

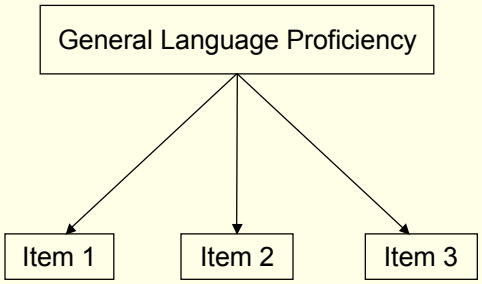
roemhild@bigred.unl.edu

Background: Language Proficiency Studies

- Language proficiency construct:
  - Unidimensional construct (Oller, 1979)
  - Multidimensional construct (Morgan & Mazzeo, 1988)
  - Hierarchical structure (Bachman et al, 1990; Sasaki, 1996)

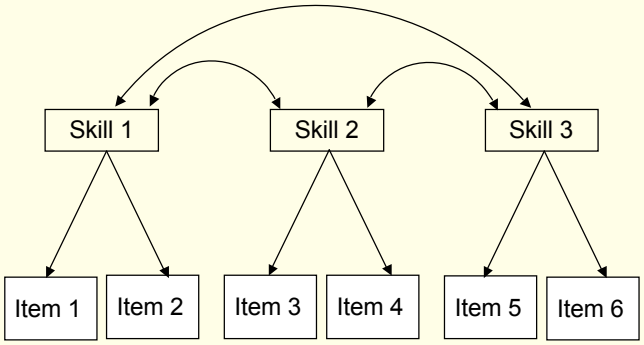
Background: Language Proficiency Studies

- Unidimensional construct (Oller, 1979)



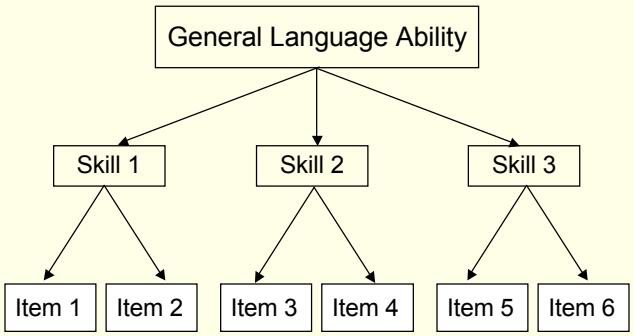
Background: Language Proficiency Studies

- Multidimensional construct (Morgan & Mazzeo, 1988)



Background: Language Proficiency Studies

- Hierarchical structure (Bachman et al, 1990; Sasaki, 1996)



Background: Learner Populations

- Population samples in dimensionality studies:
  - College students and school children
  - Little out-of-school language experience
  - Homogeneity of education level and background
- Limited English Proficient (LEP) student population in the U.S.:
  - Substantial exposure to target language outside of school
  - Heterogeneity of education level and background

Background: Invariance Studies

- Lack of invariance of factorial structure across different proficiency level groups
- Lack of invariance of factorial structure due to out-of-school language experience (Morgan & Mazzeo, 1988)
- Differential rates of acquisition of conversational fluency and grade-appropriate academic literacy skills in immigrant children (Cummins, 1980)

Research Questions:

- What is the factorial structure of an English Language Proficiency test administered to LEP students in the U.S. ?
  - Single factor model
  - Uncorrelated multiple factors model
  - Correlated multiple factors model
  - Hierarchical factor structure model
- Are the structural relationships of the test the same across different groups of LEP students?
  - Language Proficiency Level
  - Language Learning Context

Instrument: Stanford English Language Proficiency Test (SELP)

- Form B of the augmented version of the SELP used in the 2005 spring administration of the Virginia state testing program for LEP students

Table 1. VASELP Composition

Test levels		*VS0*	*VS-E1*	*VS-E2*	*VS-M*	*VS-H*
Grade levels per Test		K-1	2	3-5	6-8	9-12
Reliability (α)			0.95	0.96	0.96	0.96
Subtests	*Item Type*		*Number of Items*			
Listening	Multiple choice		20	20	20	20
Reading			20	36	44	52
Writing Conventions			20	20	24	24
Direct Writing (0-4)	Performance		2	2	2	2
Speaking (0-2, 0-4)			16	16	16	16

Methods: Data Preparation

- Creating testlets to avoid statistical problems with non-normally distributed item scores
  - by common activity type (Speaking)
  - by similar means and standard deviations (Reading, Listening, Writing Conventions)
  - Four items per subtest

Listening:	5 testlets
Reading:	5-13 testlets
Writing Conventions:	5-6 testlets
Direct Writing:	2 writing prompts
Speaking:	3 testlets, 1 story telling item
Total:	21-29 indicator variables

Sample: Student Groups and Sample Size

- New Student Category: students who have been enrolled in a U.S. school for less than one school year
  - Less out-of-school language experience
  - Lower language proficiency

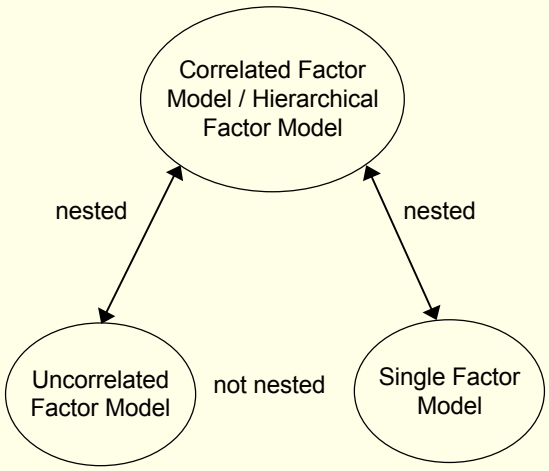
Table 2. ANOVA Results for Proficiency Level Differences

	Cont. Students	New Students	*Listening*	*Reading*	*WC*	*DW*	*Speaking*
					F (p)		
VS-E1	6388	57	35.96 (<.001)	109.4 (<.001)	76.68 (<.001)	82.67 (<.001)	77.37 (<.001)
VS-E2	15877	572	422.35 (<.001)	861.95 (<.001)	683.18 (<.001)	455.57 (<.001)	571.01 (<.001)
VS-M	9132	303	0.23 (p=0.63)	0.01 (p=0.945)	0.83 (p=0.363)	0.03 (p=0.861)	0.55 (p=0.458)
VS-H	8765	126	60.46 (<.001)	47.89 (<.001)	63.81 (<.001)	100.85 (<.001)	34.68 (<.001)

Methods: Factorial Structure of VASELP tests

- Structural equation modeling to test hypothesized language proficiency models
  - Exploratory Factor Analysis to guide determination of configural factor structure for each group
  - Establish a baseline model with multiple correlated factors for each group
  - Test and compare hypothesized models to baseline model in single-group analyses

Methods: Model Comparison



Methods: Invariance tests

- Multiple group analyses:
  - Metric invariance: Testing equivalence of relationships between factors and indicators
  - Structural invariance: Testing equivalence of factor relationships

EFA Results: Principal Components of VASELP

Table 3. Extracted Eigenvalues

		Eigenvalues (% Variance explained)			
		1	*2*	*3*	*4*
VS-E1	new	14.79 (70.5)	1.06 (5.0)	0.92 (4.4)	0.63 (3.0)
	cont.	7.92 (37.7)	4.18 (19.9)	0.97 (4.6)	0.82 (3.9)
VS-E2	new	16.11 (64.4)	1.11 (4.4)	0.75 (3.0)	0.70 (2.8)
	cont.	9.50 (38.0)	5.13 (20.5)	1.25 (5.0)	0.93 (3.7)
VS-H	new	14.04 (46.8)	1.71 (5.7)	1.33 (4.4)	1.09 (3.6)
	cont.	12.67 (42.2)	1.75 (5.8)	1.31 (4.4)	1.25 (4.2)

EFA Results: Configural Factor Structure of VS-E1

Table 4. VS-E1 Factor Loadings*

	Cont. Students		*New Students*	
Subtests	*Factor 1*	*Factor 2*	*Factor 1*	*Factor 2*
Direct Writing	-	0.69, 0.69	0.80, 0.75	0.24
Writing Conv.	-	0.60-0.71	0.77-0.93	-
Reading	-	0.47-0.73	0.70-0.95	-0.27
Listening	0.63-0.95	-	0.73-0.81	-
Speaking	0.68-0.93	-	0.62-0.82	0.47-0.52
Factor Corr.	r=0.31		r=0.35	

* Factor loadings below 0.2 are excluded.

Factor Analysis:
- ML estimation
- oblique (oblimin) rotation*

Results: VS-E1 Model Fit Evaluation

Table 5. VS-E1 Fit Indices

Cont. Students	*χ2*	*χ2 –Diff.*	*CFI (>0.9)*	*RMSEA (<0.05)*	*SRMR (<0.08)*	*AIC*	*BIC*
2 corr. Factors (df=188)	11952*		0.865	0.099	0.062	317196	317196
1 general Factor (df=189)	37771*	25819*	0.568	0.176	0.201	342722	343006
2 uncorr. Factors (df=189)	12516*	564*	0.858	0.101	0.138	317468	317752
New Students							
2 corr. Factors (df=188)	362*		0.878	0.127	0.053	3177	3265
1 general Factor (df=189)	409*	47*	0.846	0.143	0.055	3222	3308
2 uncorr. Factors (df=189)	453*	91*	0.814	0.457	0.31	3266	3352

* p<0.001

Results: VS-E1 Multiple Group Analysis

- Correlated 2 Factor Model:
  - Structural non-invariance of factor relationship (χ2-diff.= 80.301, df=1, p<0.001)
  - Metric non-invariance for all factor loadings (χ2-diff.= 41.213, df=19, p=0.002)

EFA Results: Configural Factor Structure of VS-E2

Table 6. VS-E2 Factor Loadings*

	Cont. Students		New Students	
Subtests	Factor 1	Factor 2	Factor 1	Factor 2
Direct Writing	0.63, 0.65	-	0.65, 0.67	0.22, 0.23
Writing Conv.	0.60-0.66	-	0.73-0.82	-
Reading	0.63-0.77	-	0.74-0.93	-
Listening	-	0.56-0.81	0.56-0.72	0.23
Speaking	-0.21	0.89-0.98	0.39	0.45-0.92
Factor Corr.	r=0.29		r=0.77	

* Factor loadings below 0.2 are excluded.

Factor Analysis:
- ML estimation
- oblique (oblimin) rotation

Results: VS-E2 Model Fit Evaluation

Table 7. VS-E2 Fit Indices

Cont. Students:	*χ2*	*χ2 –Diff.*	*CFI (>0.9)*	*RMSEA (<0.05)*	*SRMR (<0.08)*	*AIC*	*BIC*
2 corr. Factors (df=274)	35836*		0.876	0.09	0.069	1022688	1023080
1 general Factor (df=275)	152398*	116561*	0.468	0.187	0.19	1139248	1139632
2 uncorr. Factors (df=275)	36935*	1098*	0.872	0.092	0.134	1023785	1024168
New Students:							
2 corr. Factors (df=274)	1745*		0.897	0. 097	0.045	38941	39163
1 general Factor (df=275)	2128*	383*	0.87	0.107	0.041	39322	39539
2 uncorr. Factors (df=275)	2554*	809*	0.84	0.12	0.404	39748	39965

* p<0.001

Results: VS-E2 Multiple Group Analysis

- Correlated 2 Factor Model:
  - Structural non-invariance of factor relationship (χ2-diff.= 620.106, df=1, p<0.001)
  - Metric non-invariance for all factor loadings (χ2-diff.= 1143.431, df=24, p<0.001)

EFA Results: Configural Factor Structure of VS-H

Table 8. VS-H Factor Loadings*

	Cont. Students		New Students	
Subtests	Factor 1	Factor 2	Factor 1	Factor 2
Direct Writing	0.47, 0.49	-	0.52, 0.44	0.29, 0.39
Writing Conv.	0.48-0.59	-	0.24-0.57	0.20-0.24
Reading	0.64-0.77	-	0.57-0.80	-0.24
Listening	0.41-0.48	0.21, 0.22	0.28-0.46	0.22-0.36
Speaking	-	0.65-0.84	-	0.69-0.81

* Factor loadings below 0.2 are excluded.

Factor Analysis:
- ML estimation
- oblique (oblimin) rotation

Results: VS-H Model Fit Evaluation

Table 9. VS-H Fit Indices

Cont. Students:	*χ2*	*χ2 –Diff.*	*CFI (>0.9)*	*RMSEA (<0.05)*	*SRMR (<0.08)*	*AIC*	*BIC*
2 corr. Factors (df=404)	17983*		0.875	0.07	0.047	694900	695332
1 general Factor (df=405)	30216*	12233*	0.788	0.092	0.058	707131	707557
2 uncorr. Factors (df=405)	23191*	5208*	0.838	0.08	0.186	700106	700531
New Students:							
2 corr. Factors (df=404)	694*		0.876	0. 076	0.06	10167	10340
1 general Factor (df=405)	825*	131*	0.82	0.091	0.064	10297	10467
2 uncorr. Factors (df=405)	823*	129*	0.821	0.091	0.232	10295	10465

* p<0.001

Results: VS-H Multiple Group Analysis

- Correlated 2 Factor Model:
  - Structural non-invariance of factor relationship (χ2-diff.= 8.137, df=1, p<0.005)
  - Metric non-invariance for all factor loadings (χ2-diff.= 66.699, df=29, p<0.001)

Summary and Discussion of Results:

- Non-invariance of factor correlations and factor loadings between the two groups at each analyzed test level
- Language proficiency of students with more language experience in the US tends to be multidimensional whereas new students' language proficiency is more unidimensional
- Students classified LEP in the U.S. at an older age (High school) seem to exhibit a less differentiated factor structure than younger students

Summary and Discussion of Results:

- Students in the Elementary grades learn English while also developing academic skills such as reading and writing
- LEP High school students are more likely to have already developed academic skills in their first language, which may influence how they learning to read and write in a second language → their language proficiency is less differentiated

Table 9. Factor Correlations in 2 Factor Model

	Continuing	*New*
VS-E1	r=0.315	r=0.93
VS-E2	r=0.275	r=0.916
VS-H	r=0.759	r=0.857

Limitations:

- Language background is a potential confounding variable for non-invariance findings
- Differences in the factorial structure between test levels may be artifacts of differences in the construct representation of the tests
- Small Sample Size

Thank you!

roemhild@bigred.unl.edu