Identifying the Correct Number of Classes in Latent Profile Analysis: The Impact of Sample Size, Profile Distribution, and Model Specification

Sara K. Johnson, Ph.D.

Modern Modeling Methods

May 20, 2014

Institute for Applied Research
in Youth Development

LPAs, everywhere…

Parent-child relations (Bowers et al., 2014)

Neighborhood recreation environments (Norman et al., 2010)

Assessment practices of child clinicians (Cook, 2013)

Acculturation (Fox et al., 2013)

But, How Should We Choose the # of Profiles?

Most empirically-based advice is from studies of LCAs (Nylund et al., 2007)

But LPAs are a bit different (e.g., assumption of local independence not necessary)

Guidance on Class/Profile Enumeration

* **Nylund et al. (2007)**
  * **BIC** **(especially with small samples)**
  * **Bootstrapped Likelihood Ratio Test**
* **Lubke & Muthen (2007)**
  * **Adjusted Lo-Mendell-Rubin Likelihood Ratio Test** (especially with high profile separation)
  * **Retained conditional independence assumption**
* **Peugh & Fan (2013)**

Peugh & Fan (2013)

* **Conditions**
  * **Number of indicators**
  * **Sample Size**
  * **Profile Separation**

* **Findings**
  * **Consistently low rates of identification**
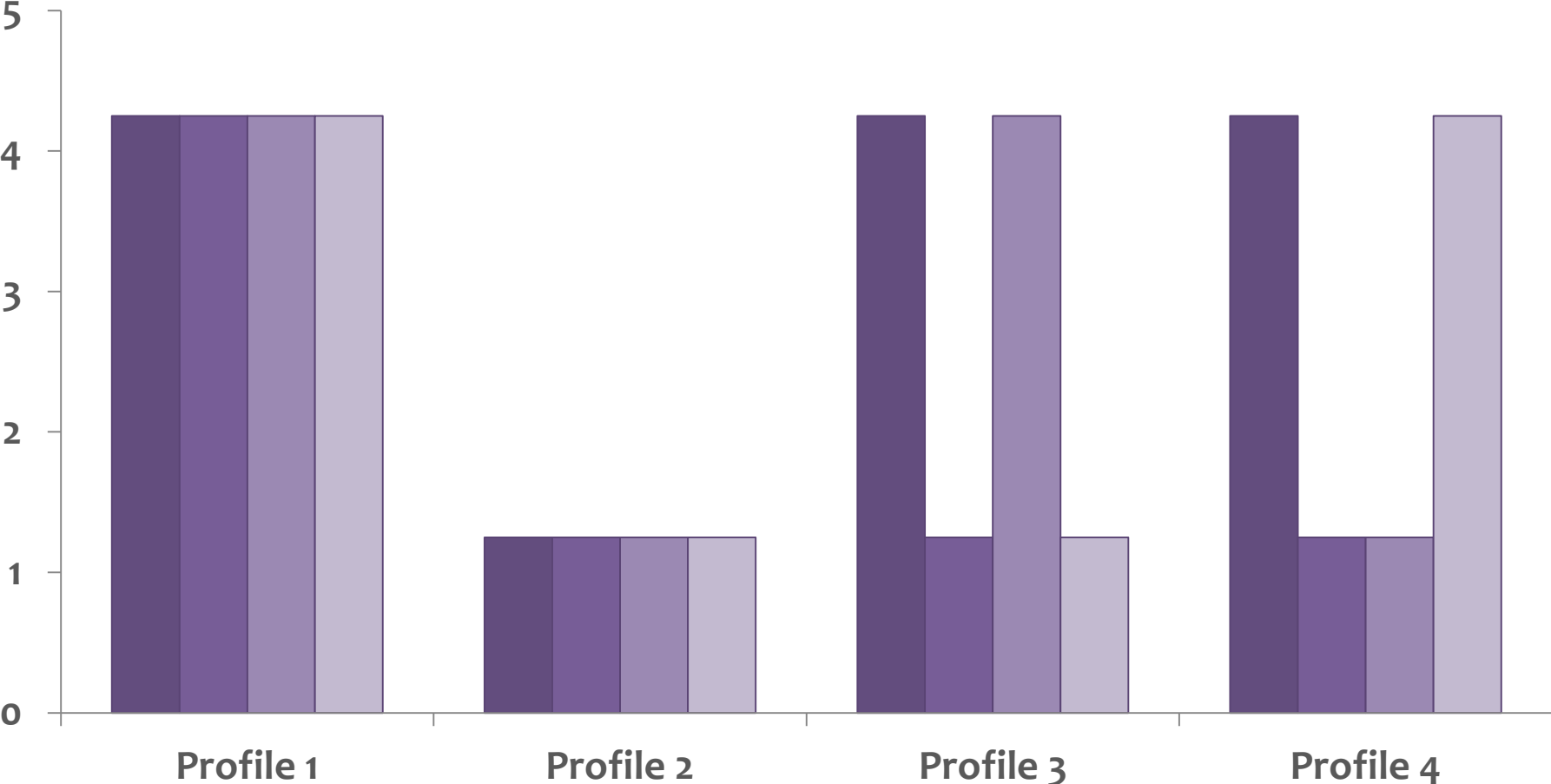 * **Most (except Entropy) overextracted**

* **Limitations**
  * **Parallel Profiles**
  * **Equal Profile Distributions**
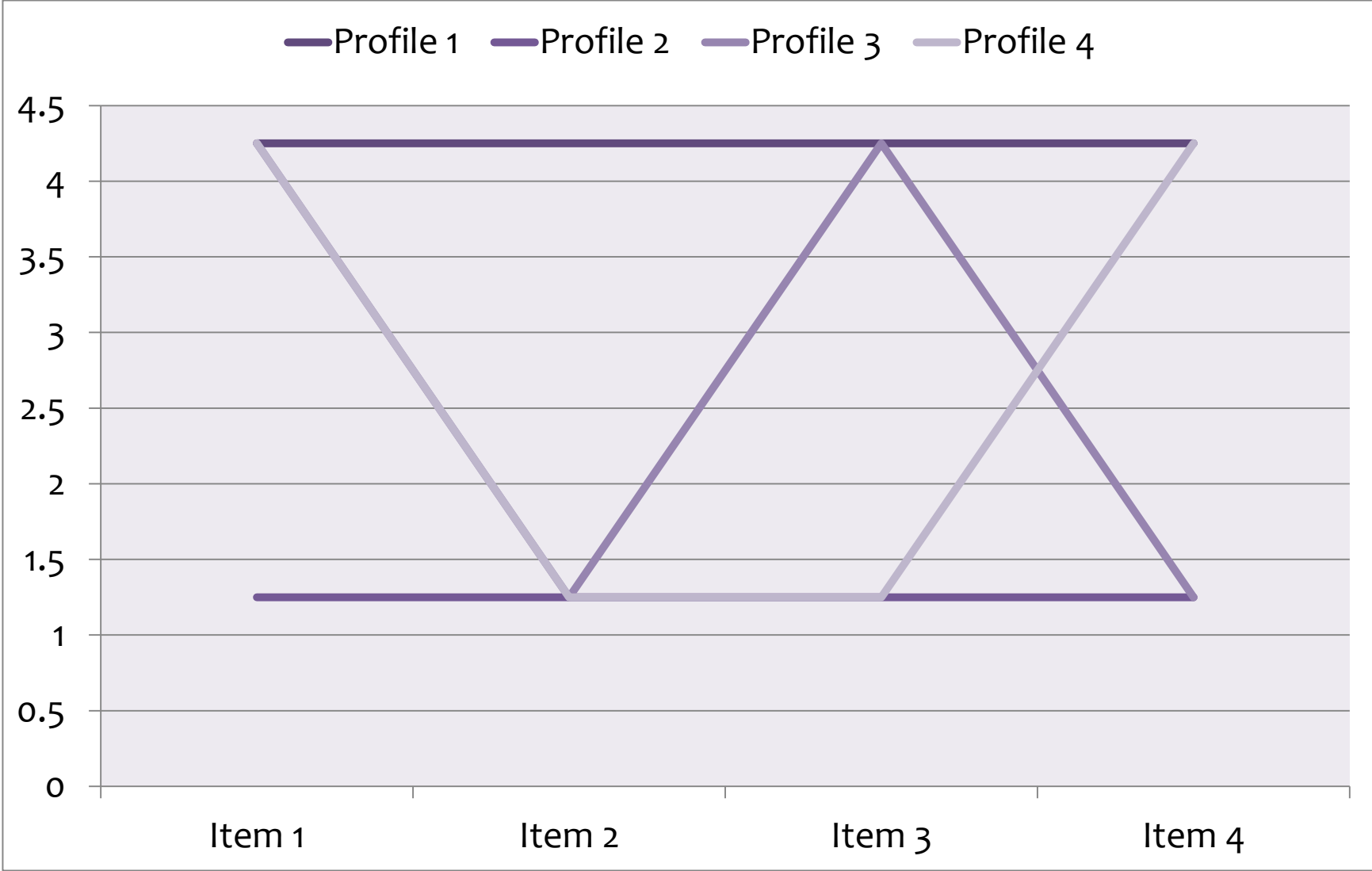 * **Two Model Specifications**

Purpose

To evaluate the performance of various fit indices in identifying the correct number of profiles in Latent Profile Analysis (LPA) depending on various simulation conditions
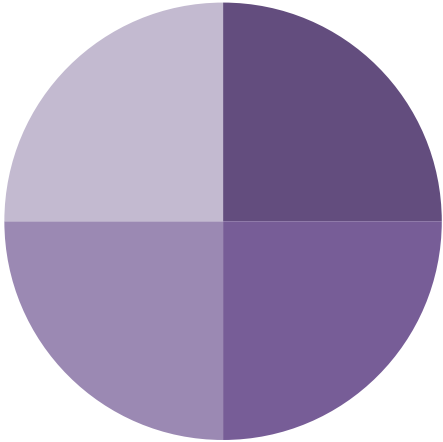
Profile Shape

Profile Shape, Take 2

Simulation Conditions

* **Profile Distributions**

* **Sample Size**

* **Model Specification**

Profile Distributions

Equal Distribution

Large First Class
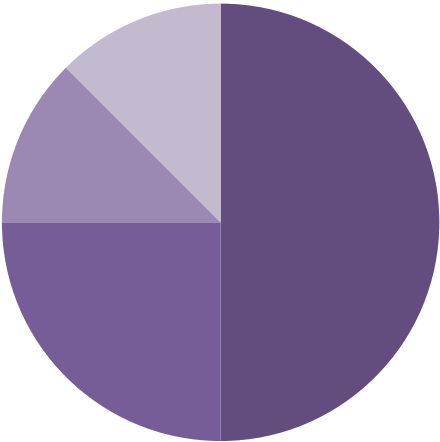
Very Small ("Tiny") Class

Sample Sizes

* **300**

* **600**
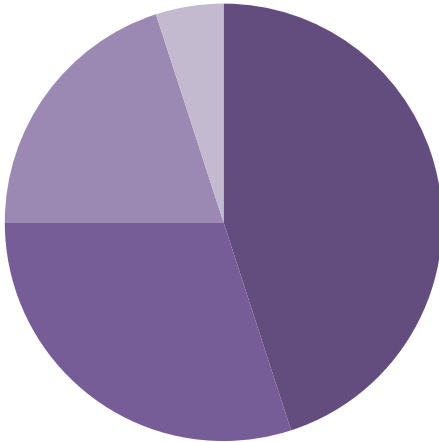
* **900**

* **1200**

* **1500**

* **3000**

Sample Sizes

* **300**
* **600**
* **900**
* **1200**
* **1500**
* **3000**

Methods of Analysis

* **No Correlations** (Conditional Independence)

* **Correlations Equal across Profiles**

* **Correlations Unequal across Profiles**

Methods of Analysis

* **No Correlations** (Conditional Independence)

* **Correlations Equal across Profiles**

* **Correlations Unequal across Profiles**

Results:
Equal Profile Distributions

Equal Profile Distributions (BIC)

Profiles Chosen (% of Reps)

Sample Size & Specification		2	3	4	5
N=300	Equal Correlations	0	0	10	90
N=300	No Correlations	0	0	0	100
N=600	Equal Correlations	0	0	0	100
N=600	No Correlations	0	0	0	100

Equal Profile Distributions: AIC

		Profiles Chosen (% of Reps)			
Sample Size & Specification		2	3	4	5
N=300	Equal Correlations	0	0	0	100
N=300	No Correlations	0	0	0	100
N=600	Equal Correlations	0	0	0	100
N=600	No Correlations	0	0	0	100

Equal Profile Distributions: SABIC

Sample Size & Specification		Profiles Chosen (% of Reps)			
		2	3	4	5
N=300	Equal Correlations	0	0	0	100
N=300	No Correlations	0	0	0	100
N=600	Equal Correlations	0	0	0	100
N=600	No Correlations	0	0	0	100

Equal Profile Distributions: Entropy

Sample Size & Specification		Profiles Chosen (% of Reps)			
		2	3	4	5
N=300	Equal Correlations	**54.5**	29.5	13.5	2.5
N=300	No Correlations	15.5	0	**79.0**	5.5
N=600	Equal Correlations	**56.6**	31.0	12.5	0
N=600	No Correlations	11.0	0	**89.0**	0

Equal Profile Distributions: LMR

Sample Size & Specification	Profiles Chosen (% of Reps)			
	1	2	3	4
N=300 Equal Correlations	0	0	11.0	40.0
N=300 No Correlations	0	0	0.5	39.5
N=600 Equal Correlations	0	0	6.5	13.0
N=600 No Correlations	0	0	0	21.5

Results:
Large First Class

Large First Class: BIC

Sample Size & Specification		Profiles Chosen (% of Reps)			
		2	3	4	5
N=300	Equal Correlations	0	0	44	56
N=300	No Correlations	0	0	0	100
N=600	Equal Correlations	0	0	5	95
N=600	No Correlations	0	0	0	100

Large First Class: AIC

	Profiles Chosen (% of Reps)			
Sample Size & Specification	2	3	4	5
N=300 Equal Correlations	0	0	0	100
N=300 No Correlations	0	0	0	100
N=600 Equal Correlations	0	0	0	100
N=600 No Correlations	0	0	0	100

Large First Class: SABIC

Sample Size & Specification	Profiles Chosen (% of Reps)			
	2	3	4	5
N=300 Equal Correlations	0	0	0	100
N=300 No Correlations	0	0	0	100
N=600 Equal Correlations	0	0	0	100
N=600 No Correlations	0	0	0	100

Large First Class: Entropy

Sample Size & Specification		Profiles Chosen (% of Reps)			
		2	3	4	5
N=300	Equal Correlations	**80.5**	5.5	3.5	10.5
N=300	No Correlations	18.0	3.5	**67.0**	11.5
N=600	Equal Correlations	**88.5**	2.0	4.5	5.0
N=600	No Correlations	11.5	1.5	**83.5**	3.5

Large First Class: LMR

Profiles Chosen (% of Reps)

Sample Size & Specification		1	2	3	4
N=300	Equal Correlations	1.0	0.5	2.0	**46.0**
N=300	No Correlations	0.5	11.5	26.5	**31.5**
N=600	Equal Correlations	0.5	0.5	0	**27.5**
N=600	No Correlations	0	1.5	9.5	**25.5**

Results:
Tiny Class

Tiny Class: BIC

Sample Size & Specification		Profiles Chosen (% of Reps)			
		2	3	4	5
N=300	Equal Correlations	0	0	0	100
N=300	No Correlations	0	0	0	100
N=600	Equal Correlations	0	0	6.5	93.5
N=600	No Correlations	0	0	0	100

Tiny Class: AIC

Sample Size & Specification		Profiles Chosen (% of Reps)			
		2	3	4	5
N=300	Equal Correlations	0	0	0.5	95.5
N=300	No Correlations	0	0	0	100
N=600	Equal Correlations	0	0	0	100
N=600	No Correlations	0	0	0	100

Tiny Class: SABIC

Profiles Chosen (% of Reps)

Sample Size & Specification		2	3	4	5
N=300	Equal Correlations	0	0	0	100
N=300	No Correlations	0	0	0	100
N=600	Equal Correlations	0	0	0	100
N=600	No Correlations	0	0	0	100

Tiny Class: Entropy

Sample Size & Specification		Profiles Chosen (% of Reps)			
		2	3	4	5
N=300	Equal Correlations	**64.0**	5.0	17.0	14.0
N=300	No Correlations	**47.5**	17.0	10.5	25.0
N=600	Equal Correlations	**72.5**	1.5	21.5	4.5
N=600	No Correlations	**51.0**	24.5	2.0	22.5

Tiny Class: LMR

Sample Size & Specification		Profiles Chosen (% of Reps)			
		1	2	3	4
N=300	Equal Correlations	0	0	0	40.0
N=300	No Correlations	0.5	15.5	18.5	36.0
N=600	Equal Correlations	0	0.5	.5	28.0
N=600	No Correlations	0	11.5	6.0	31.0

Discussion

Discussion

* ## Information Criteria
  * ### Consistently Over Extract Profiles
* ## Entropy
  * ### Under Extracts
    * Equal Distribution – Equal Correlation
    * Large First Class – Equal Correlation
    * Tiny Class – Equal AND No Correlation
  * ### Primarily On Target
    * Equal Distribution – No Correlations
    * Large First Class – No Correlations

Discussion

* **Lo-Mendell-Rubin**
  * **Getting there!**
  * **Works best with small sample sizes**

Discussion

* **Finishing the Simulation!**
  * **Bigger sample sizes**
  * **BLRTs**
  * **Unequal correlations**
* **Future Directions**
  * **Profile Composition**
    * **Lower rates of separation**
    * **Different profile shapes**
  * **More indicators**

Thank You!