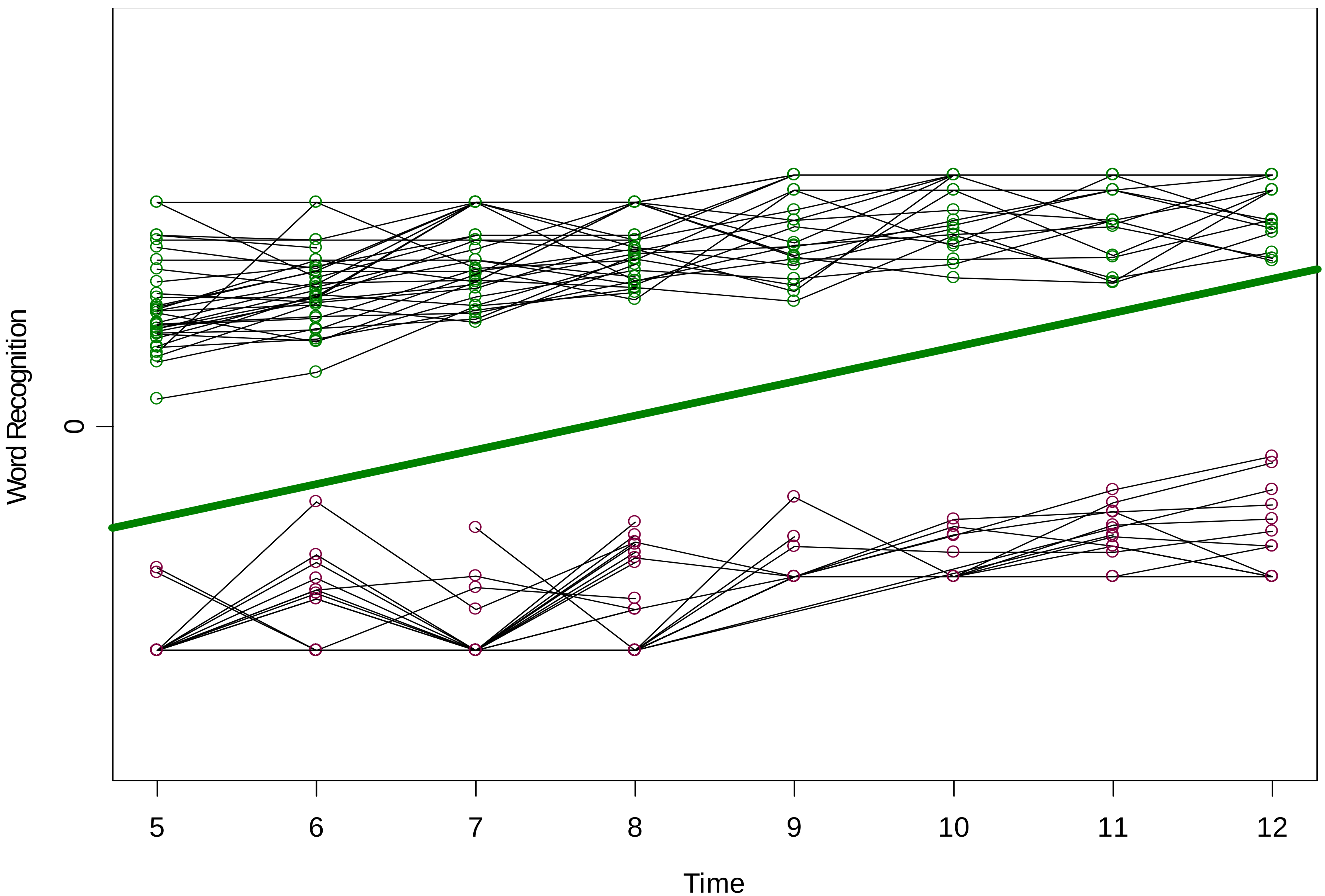
APPLICATION OF GROWTH MIXTURE MODELS FOR EARLY DETECTION

Christy K. Boscardin, Ph.D.

CAPS Methods Core Seminar

October 30, 2009

- Boscardin, C., Muthén, B., Francis, D. & Baker, E. (2008). *Early identification of reading difficulties using heterogeneous developmental trajectories*. Journal of Educational Psychology, 100, 192-208.



Word Recognition
0
5
6
7
8
9
10
11
12
Time

Advantages of Growth Mixture

- Provide more flexibility of assumptions about the population
- Used to identify heterogeneous subgroups in the population
- Used to identify & cluster individuals with similar developmental profiles/trajectories
- Useful for determining the appropriate intervention based on the factors associated with each subgroup

Overview

- Reading Development
- Longitudinal Data Analysis
- Introduction to Growth Mixture Modeling
- Early Reading Assessment Study

Background on Reading

- Reading Disability accounts for 80% of Learning Disability
- Term Learning Disability Coined by Kirk & Bateman (1962)
  - Unexpected Difficulties in Language, Learning, Reading
- Until recently – IQ discrepancy method used for diagnosis
- Discrepancy between ability (as measured by IQ) and reading achievement (as measured by standardized state test)

Problems with IQ Discrepancy Method

- 2 measures involved: IQ and standardized achievement test
- Problems with IQ test:
  - IQ includes linguistic components
  - Validity of IQ tests have been questioned
- Students with reading disability diagnosed in 2^{nd} grade
  - No standardized achievement scores available
  - Require oral assessment rather than traditional testing format
  - Unnecessarily delay identification

Background on Reading Cont.

- NICHD study: approximately 75% of children will continue to have reading difficulties in later grades diagnosed after 2nd grade
- Early intervention is key to successful remediation
- Studies with intervention starting in kindergarten shown success
- Intervention most effective prior to overt manifestation

Current Reading Development Movement

- Focus shifted to early identification of student at-risk
- Require examination of predictors & precursors
- Limited empirical studies on reading development over time
- **Hypothesis: Reading development may include subtypes rather than one normative trajectory**
- Longitudinal analysis helps identify best time to intervene

Research Questions

- Are there distinct subtypes of reading developmental trajectories?
- Are students at-risk for reading difficulties look qualitatively different compared to students with normal reading development?
- What are the factors related to being at-risk?

Growth Modeling

- Typical approach for looking at development over time: Conventional Growth Modeling
- Describes changes in pattern over time using repeated measures
- Typical questions:
- Are there individual differences in reading development over time?
- Do boys and girls develop differently?

Growth Modeling (Latent Variable Framework)

With individual repeated measures, you can estimate underlying developmental trajectory:

$$\mathbf{Y}_{ijt} = \eta_{ij1} + (t - T)\,\eta_{ij2} + \varepsilon_{ijt}\ ;\ t = 1, 2, \ldots, T.$$

Where for example

$\mathbf{Y}_{ijt}$ = *j*th reading skill for *i*th individual at time *t*

η_{ij1} = initial status of reading ability

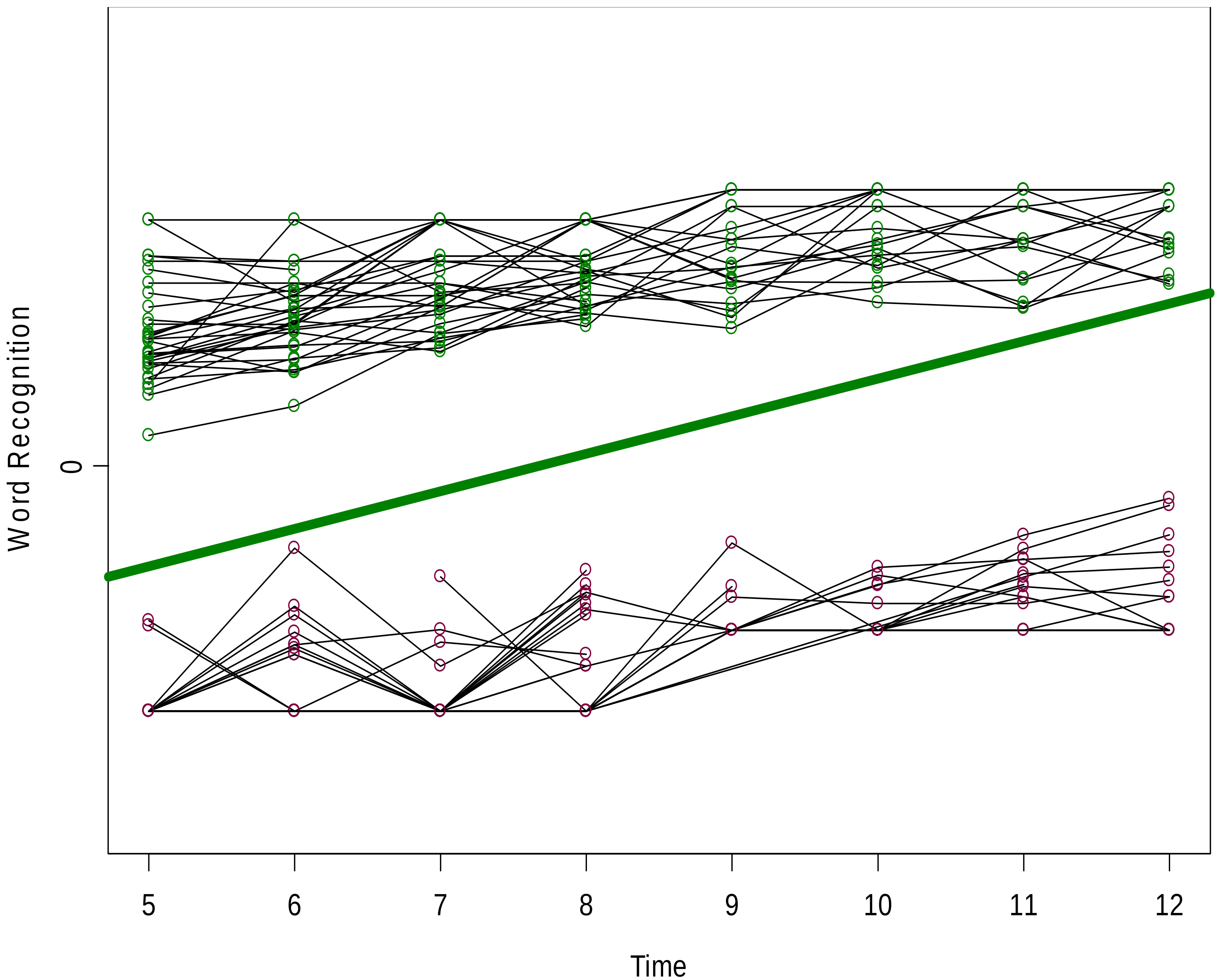
η_{ij2} = rate of improvement or change in reading ability

ε_{ijt} = error associated with the model

Individual growth parameters are assumed to come from a single population: one normative slope & one normative intercept

Growth Mixture Modeling

- Technical advantages over conventional growth modeling
- Allow more flexibility in model specifications
- Allows for heterogeneity of subgroups – more than one population for individual specific parameters
- Advantages
  - More precise estimate of growth trajectory
  - Identify variations in developmental profiles
  - Vary the effect of covariates (factors) depending on the developmental profile
- Reading development suggest subgroups
  - Student with no response to instruction may be etiologically different from students who do respond to instruction



Word Recognition
0
5
6
7
8
9
10
11
12
Time

Growth Mixture Modeling

- Here's the difference:
- Allowing for heterogeneity of subgroups with latent categorical variable
- Individuals are allowed to be in one of K latent classes (characteristically distinct developmental profiles)
- Covariates are used to both estimate growth but also to help predict class membership

Growth Mixture Modeling

$$\mathbf{Y}_{ik} = \nu_k + \Lambda_k \eta_{ik} + \mathbf{K}_k \mathbf{x}_{ik} + \varepsilon_{ik},$$
$$\eta_{ik} = \alpha_k + \mathbf{B}_k \eta_{ik} + \Gamma_k \mathbf{x}_{ik} + \zeta_{ik},$$

Where

$\mathbf{Y}_{ik}$ represents the repeated measures over fixed time points.

η_{ik} random effects

Λ_k represent the shape of the growth curves.

$\mathbf{K}_k$ represents the effects of time-varying covariates,

Γ_k represents the effects of time-invariant covariates.

α_k represents the intercepts for η for latent class k.

The residual vectors, ε_{ik} and ζ_{ik}, are assumed to have covariance matrices Θ_k and Ψ_k, respectively.

Class Membership

The multinomial logistic regression for predicting class membership with a covariate :

$$P\,(c_{ik} = 1|x_i) = \exp(\beta_{0k} + \beta_{1k}x_i)/\sum_{c=1}^{K}\exp(\beta_{0c} + \beta_{1c}x_i)$$

Estimated posterior probabilities for each individual's class membership are derived as follows:

$$p_{ik} = P\,(c_{ik} = 1|\,y_i,\, x_i) \propto P\,(c_{ik} = 1|\,x_i)\, f\,(Y_{ik}\,|\,x_i)$$

latent class membership indicators, c_{ik} , to be 1 if individual *i* belongs to class *k*, and 0 otherwise

Technical Details: Model Selection

- Identifying Number of Classes: Model Selection
  - Bayesian Information Criterion (BIC): non-nested models – smaller value
  - Entropy: precision in estimated posterior probability
- Comparison of individual estimated mean trajectories and observed mean trajectories

Early Assessment of Reading Skills (EARS) DATA

- Purpose:
  - Identify Students Most At-Risk for Reading Difficulties
  - Factors related to at-risk for reading difficulties
  - Earlier identification using precursor measures instead of waiting until 2^{nd} grade
- Sample:
  - Subset of 411 from 945 with complete data in kindergarten
    - No significant differences in measures between the subset and total population
  - 50% boys
  - 55% White, 17% African American, 16% Hispanic, 11% Asian, and 1% other
  - White & Asian (Non-Minority) vs. African American, Hispanic, & Other (Minority)

Kindergarten: Phonological Awareness

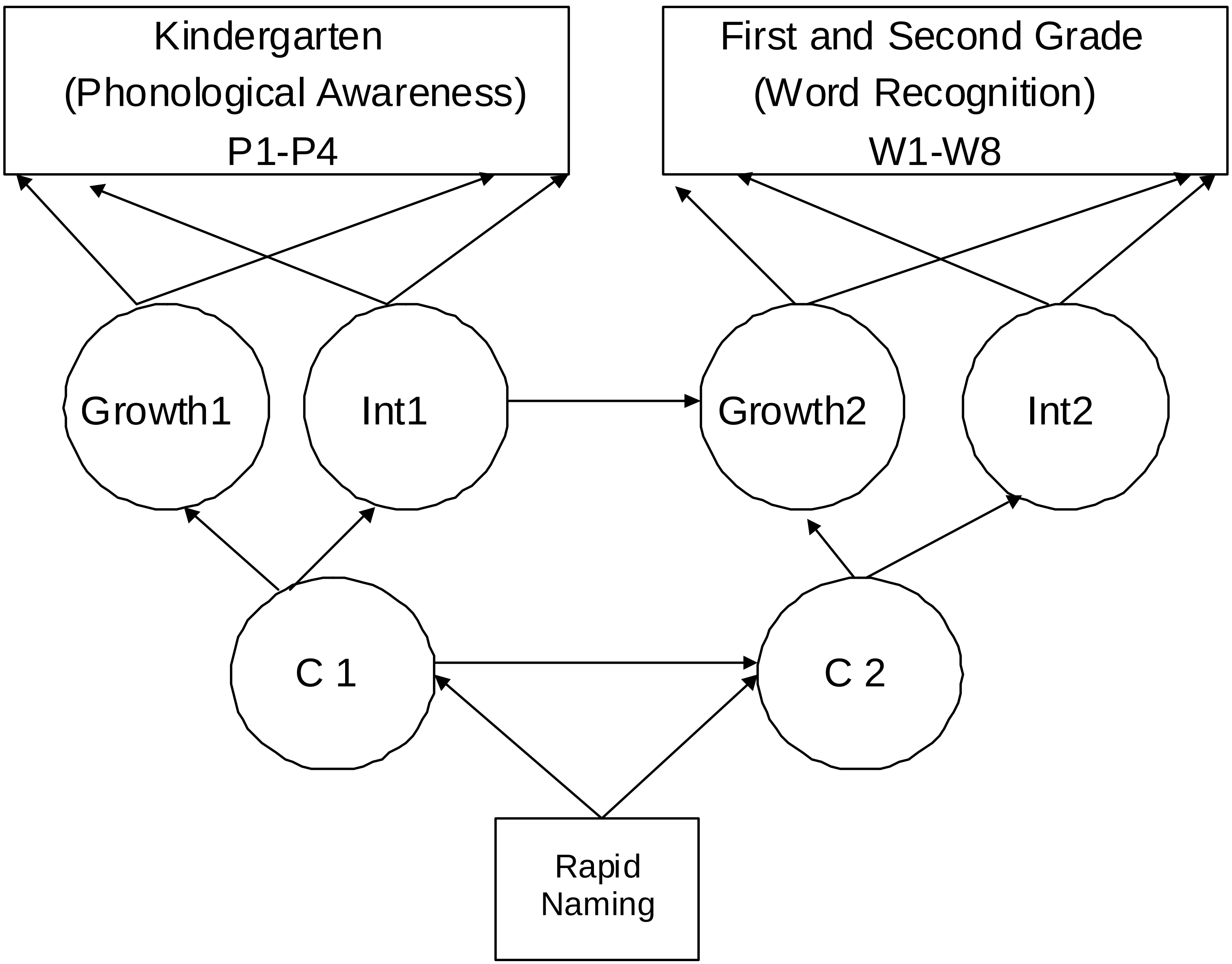
- Deficits in Phonological awareness related to reading difficulties later
- Ability to manipulate phonemes
  - cake & camp, pan & can, /s/pill & pill
- At entry to formal schooling, know 6,000 words & basic phonological awareness
- **Signs of trouble in kindergarten**: difficulty in learning to name letters and attach phonemes to letters.
- Instrument: 7 subtests including: a) phoneme segmentation, b) phoneme elision, c) sound categorization, d) first sound comparison, e) blending onset and rime, f) blending phonemes into words, and g) blending phonemes into non-words.
- Internal consistency estimates for the subtests ranged from 0.85 to 0.95
- Score based on IRT estimate
- Measured 4 times during the year: Oct, Dec, Feb, April

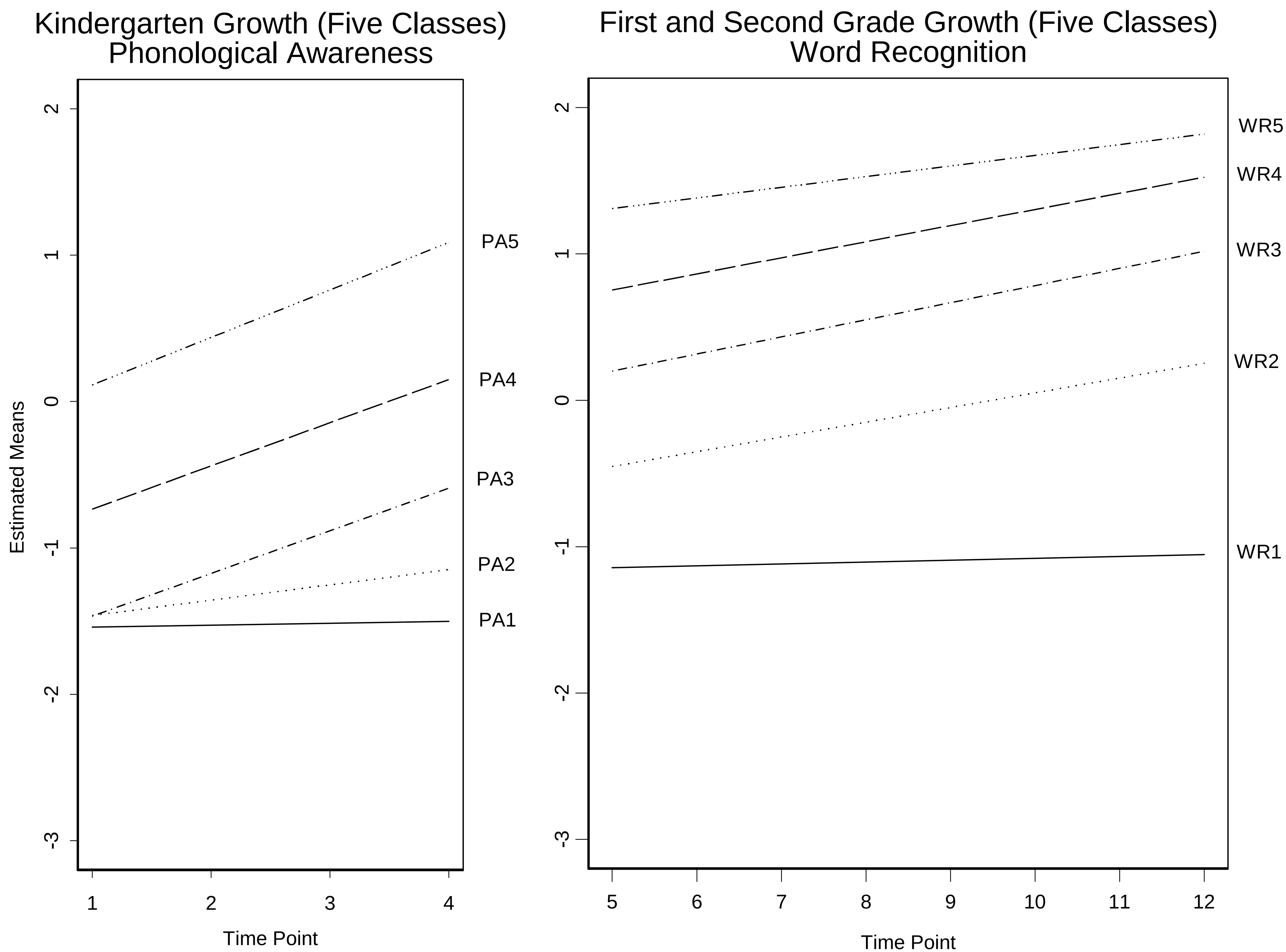
First & Second Grade: Word Recognition

- Central to reading acquisition
- Fluency and efficiency of word recognition highly correlated with reading skills in later grades
- **Signs of trouble in 1st grade:** can't recognize single words out of sentence context
- Instrument: 50 isolated words (index cards) - included 36 single-syllable, 11 two-syllable, and 3 three-syllable real words.
  - 16 words in common across the two grades
  - internal consistency estimates exceeded 0.90
  - Score based on IRT estimate
  - Measured 4 times in each grade (total of 8 times)

Covariate:
Letter Identification/Rapid Naming

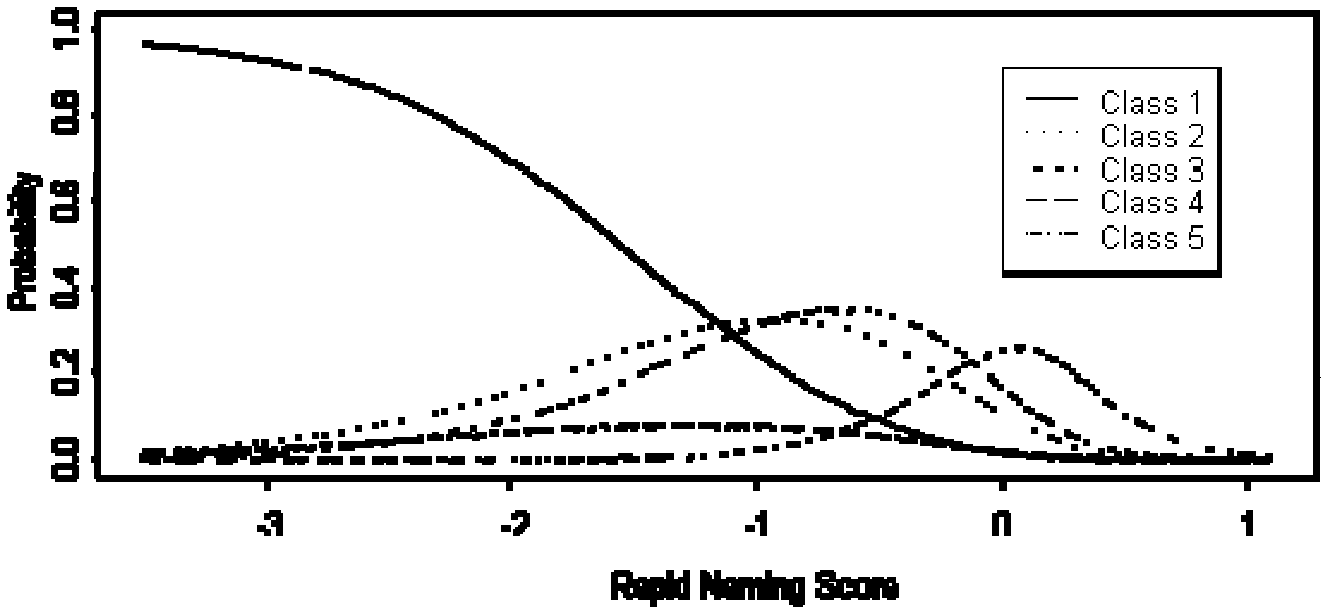
- Accuracy and Fluency of Letter Identification shown to be highly correlated with reading
- Not considered a precursor skill
- Instrument: The correct number of letter naming within 60 seconds was recorded.
  - Measured at the end of kindergarten

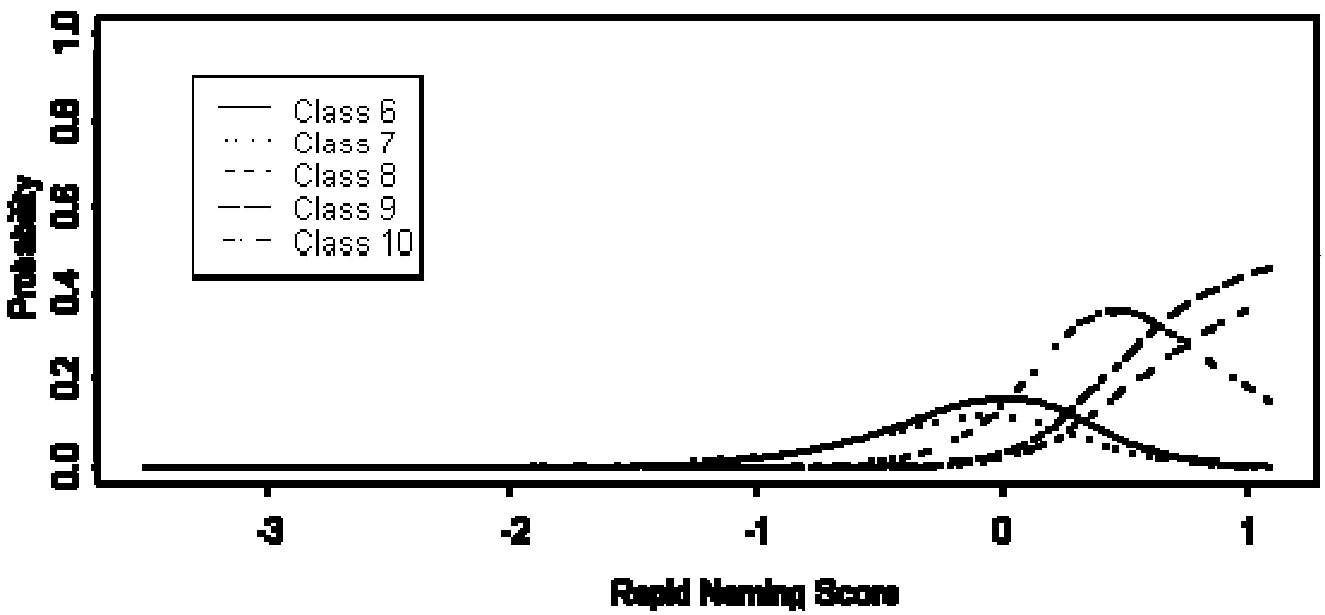

Kindergarten
(Phonological Awareness)
P1-P4
First and Second Grade
(Word Recognition)
W1-W8
Growth1
Int1
Growth2
Int2
C 1
C 2
Rapid
Naming

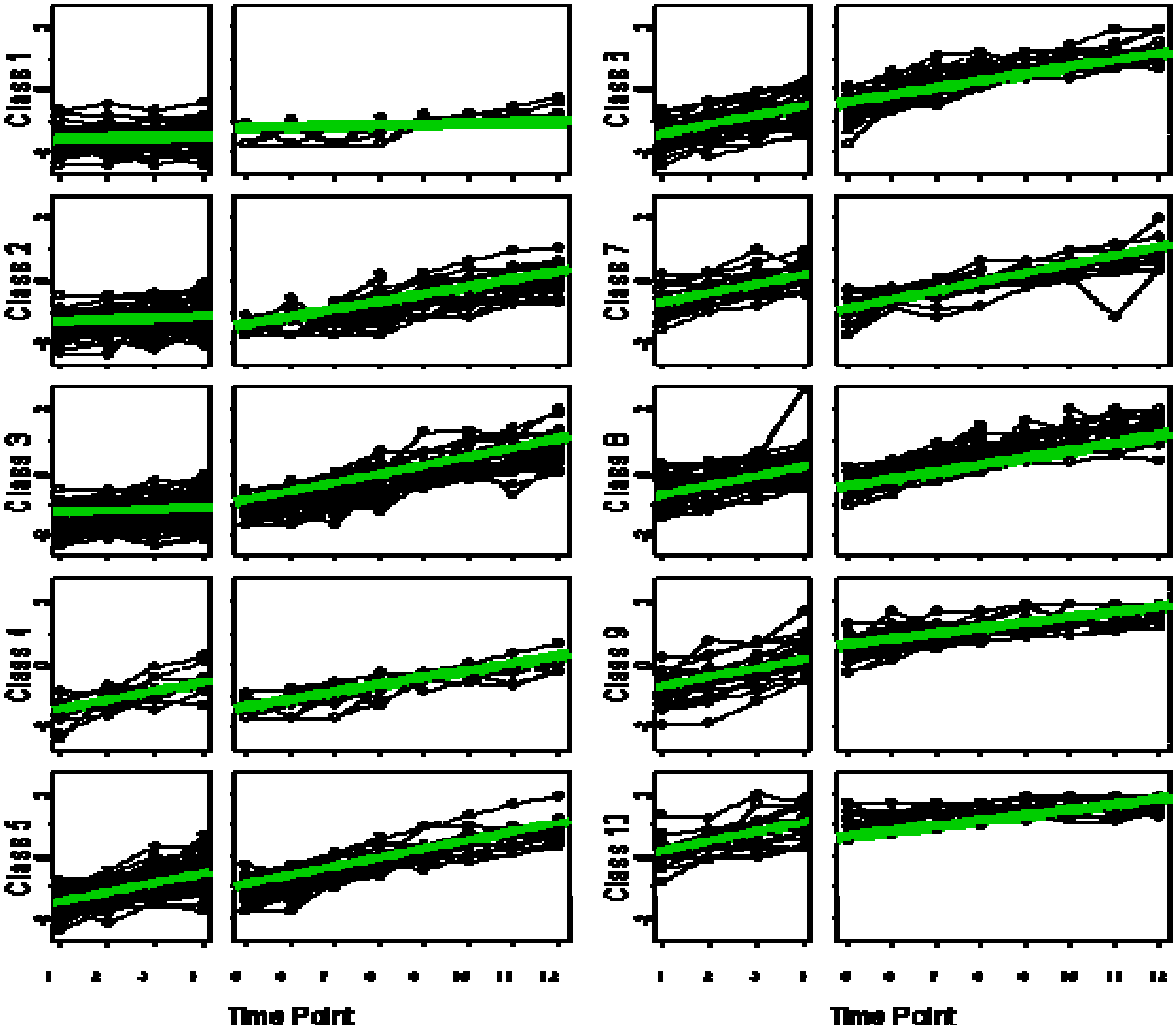


Kindergarten Growth (Five Classes)
Phonological Awareness
Estimated Means
2
1
0
-1
-2
-3
1
2
3
4
Time Point
PA5
PA4
PA3
PA2
PA1
First and Second Grade Growth (Five Classes)
Word Recognition
2
1
0
-1
-2
-3
5
6
7
8
9
10
11
12
Time Point
WR5
WR4
WR3
WR2
WR1

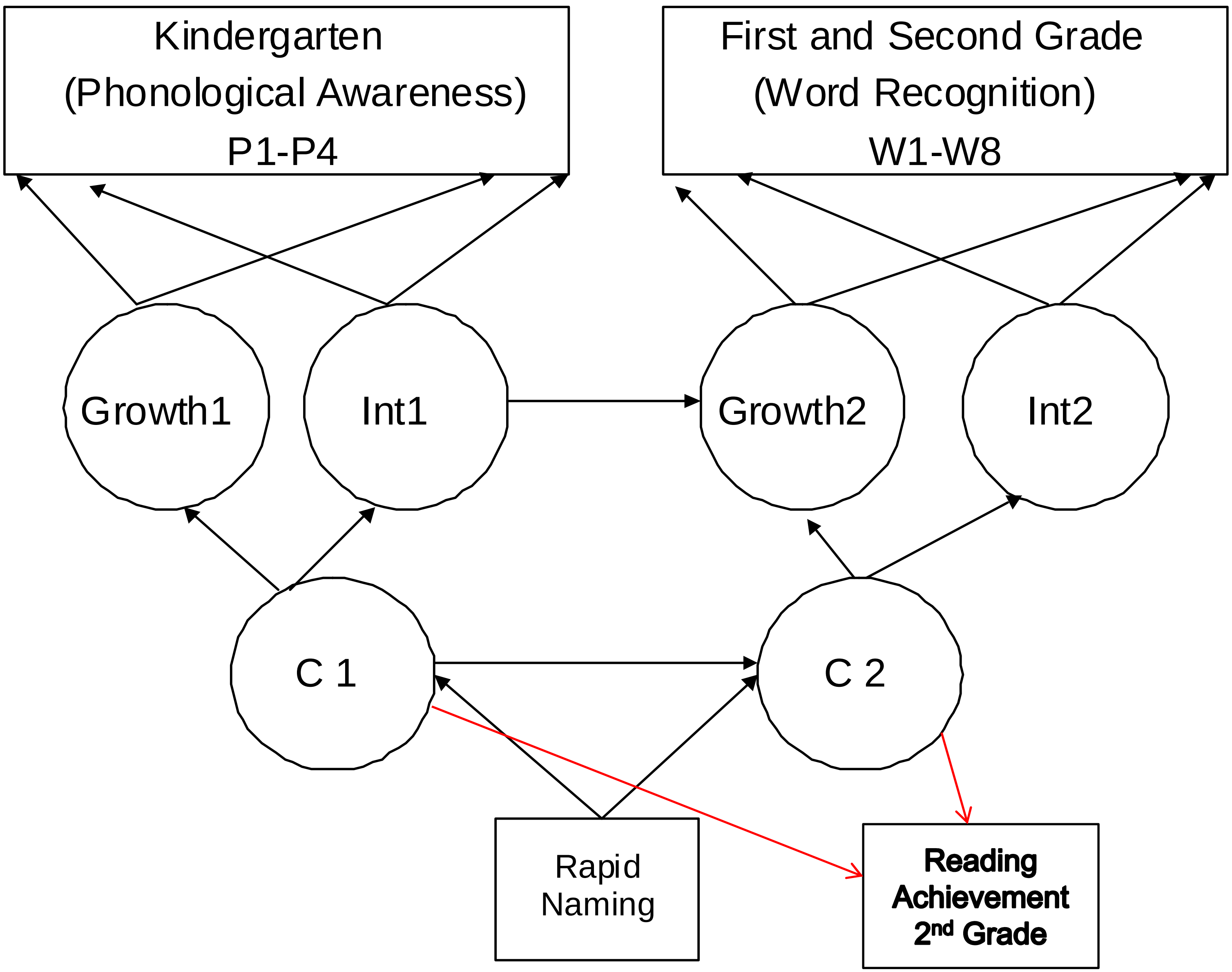
	First and Second Grade				
Kindergarten	**WR1**	**WR2**	**WR3**	**WR4**	**WR5**
PA1	**Class 1** **45 (11%)**				
PA2		**Class 2** **63 (15%)**	***Class 3*** ***77 (19%)***		
PA3		***Class 4*** ***8 (2%)***	**Class 5** **56 (14%)**	***Class 6*** ***36 (9%)***	
PA4			***Class 7*** ***20 (5%)***	**Class 8** **8 (14%)**	***Class 9*** ***30 (7%)***
PA5					**Class 10** **18 (4%)**


Class 1
Class 2
Class 3
Class 4
Class 5
Probability
1.0
0.8
0.6
0.4
0.2
0.0
-3
-2
-1
0
1
Rapid Naming Score

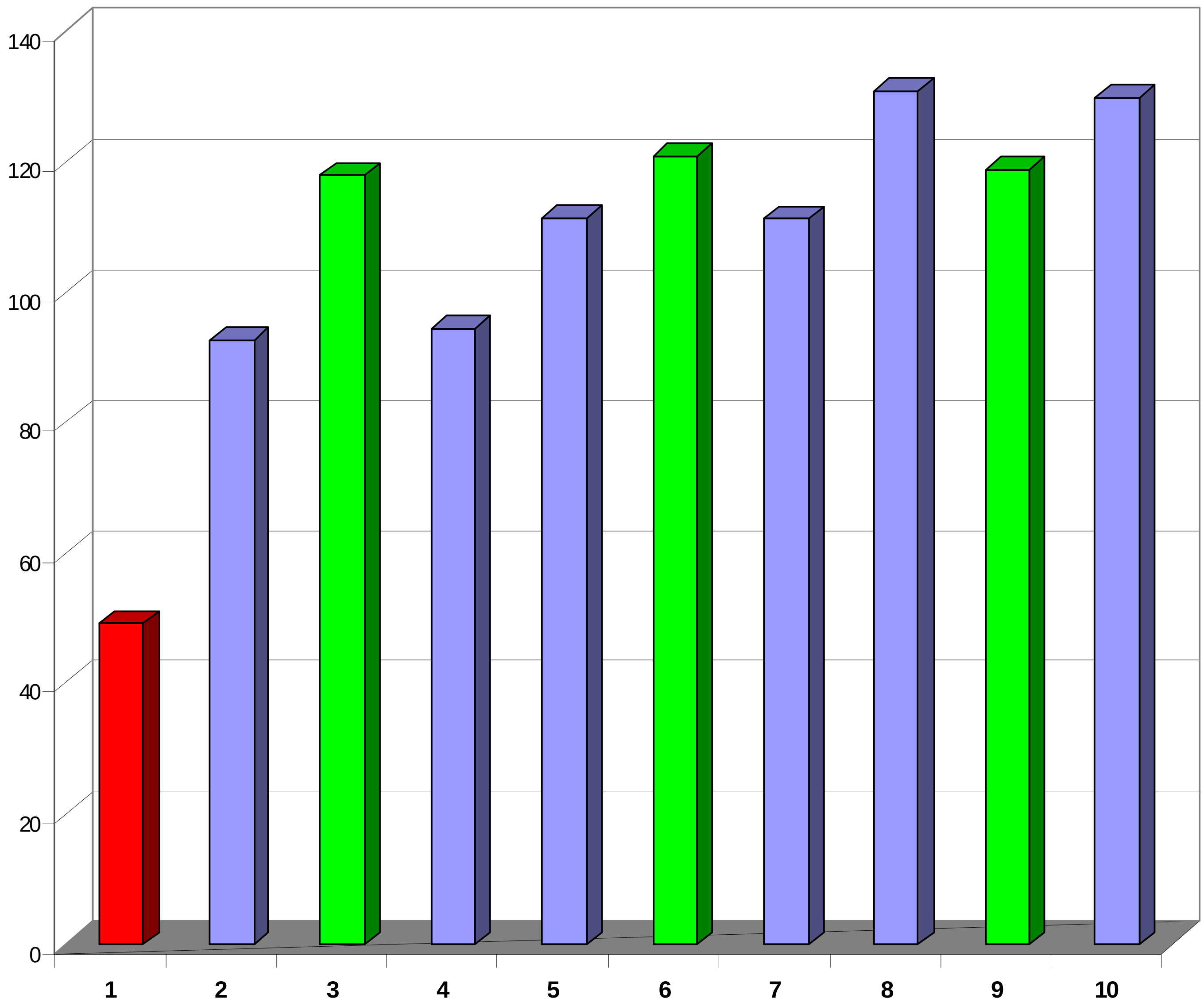

Class 6
Class 7
Class 8
Class 9
Class 10
Probability
1.0
0.8
0.6
0.4
0.2
0.0
-3
-2
-1
0
1
Rapid Naming Score



Class 1
Class 2
Class 3
Class 4
Class 5
Class 6
Class 7
Class 8
Class 9
Class 10
Time Point
Time Point


Kindergarten
(Phonological Awareness)
P1-P4
First and Second Grade
(Word Recognition)
W1-W8
Growth1
Int1
Growth2
Int2
C 1
C 2
Rapid
Naming
Reading
Achievement
2nd Grade

Reading Achievement Predicted by Class membership



Summary of Findings

- 5 Distinct Developmental Profiles in Each Grade
- 10 Distinct Developmental Profiles (including transitional groups) across 3 years
- Class 1 students: no Phonological Awareness development – no Word Recognition Development
- Rapid naming is strong predictor of class membership
- As early as kindergarten we can start to identify students who are potentially at-risk for reading difficulties
- Significantly higher proportion of minority students in class 1 compared to other classes.

Applications of Growth Mixture Modeling

- Greenbaum, P.E., Del Boca, F.K., Darkes, J., Wang, C. & Goldman, M.S. (2005). Variation in the drinking trajectories of freshman college students. Journal of Consulting and Clinical Psychology, 73, 229-238
  - Investigated whether a single growth curve adequately characterizes the variability in individual drinking trajectories. Identified 5 trajectories.

- Chen L, Eaton WW, Gallo JJ, Nestadt G: Understanding the heterogeneity of depression through the triad of symptoms, course and risk factors: A longitudinal, population-based study. J Affect Disord 2000; 59:1-11
  - There is an ongoing research effort to test if depression is a homogeneous clinical syndrome and to identify valid and useful subtypes based on the number and nature of depressive symptoms.

- Croudace, T.J., Jarvelin, M.R., Wadsworth, M.E. & Jones, P.B. (2003). Developmental typology of trajectories to nighttime bladder control: Epidemiologic application of longitudinal latent class analysis. American Journal of Epidemiology, May 1;157(9):834-42.

- Muthén, B., Brown, C.H., Masyn, K., Jo, B., Khoo, S.T., Yang, C.C., Wang, C.P., Kellam, S., Carlin, J., & Liao, J. (2002). General growth mixture modeling for randomized preventive interventions. Biostatistics, 3, 459-475.

Contact Info:

Christy K. Boscardin, PhD
christy.boscardin@ucsf.edu