

# Using Latent Class Analysis in IR: Examining Subgroups of Early Start Participants

Erika E. Baldwin, Ph.D., Senior Manager of Data Analytics  
Matthew Case, MPP, Senior Manager of Student Success Data  
Hui-Jeong Woo, Ph.D., Research Analyst

California State University Office of the Chancellor  
CAIR November 2015

## Outline

1. Background: college readiness & Early Start
2. Overview of latent class analysis (LCA)
3. How-to
4. LCA results on Early Start
5. Applying results in IR

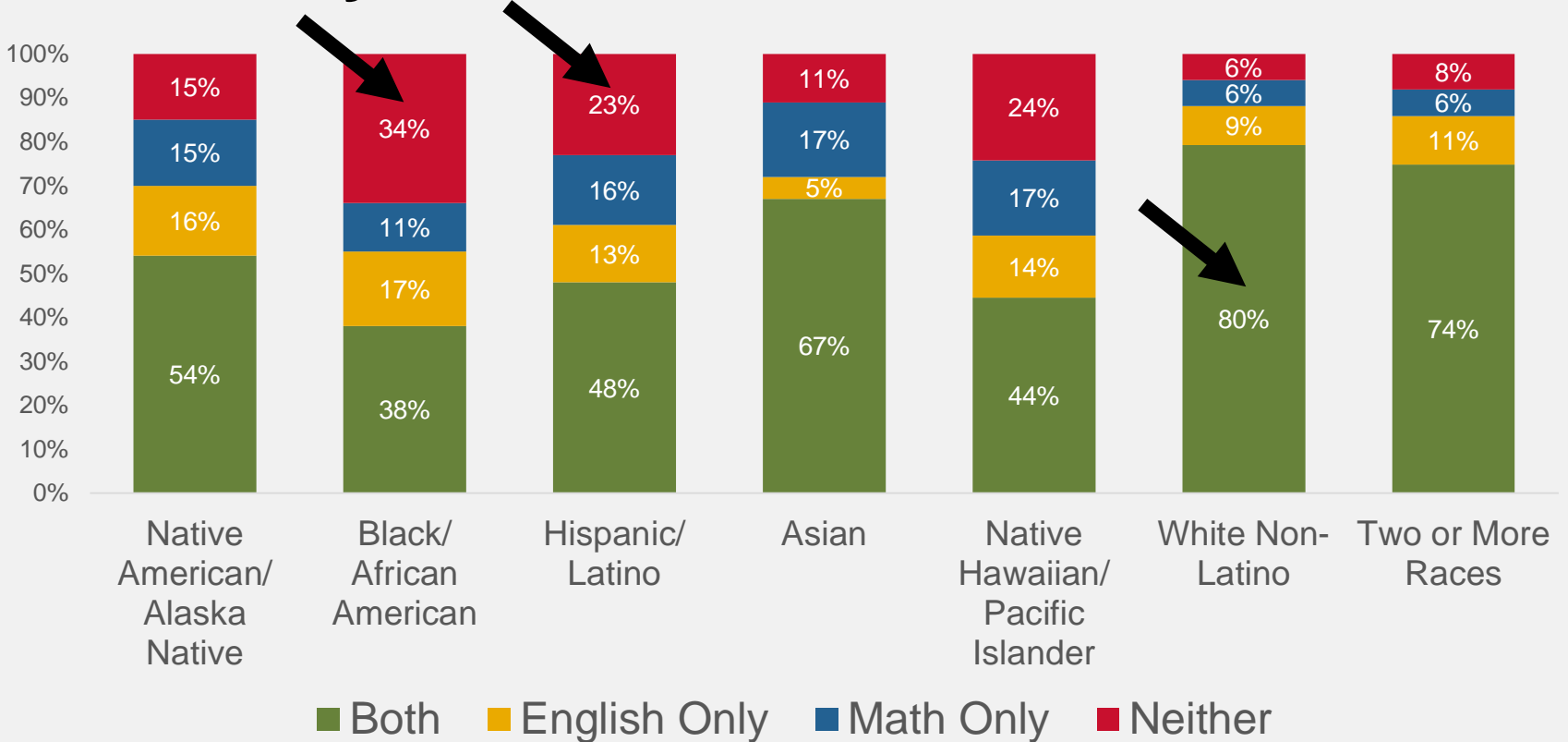
## College readiness: Fall 2014 first-time freshmen

- 59% students were college-ready in both English and math
- English
  - 72% males
  - 68% females

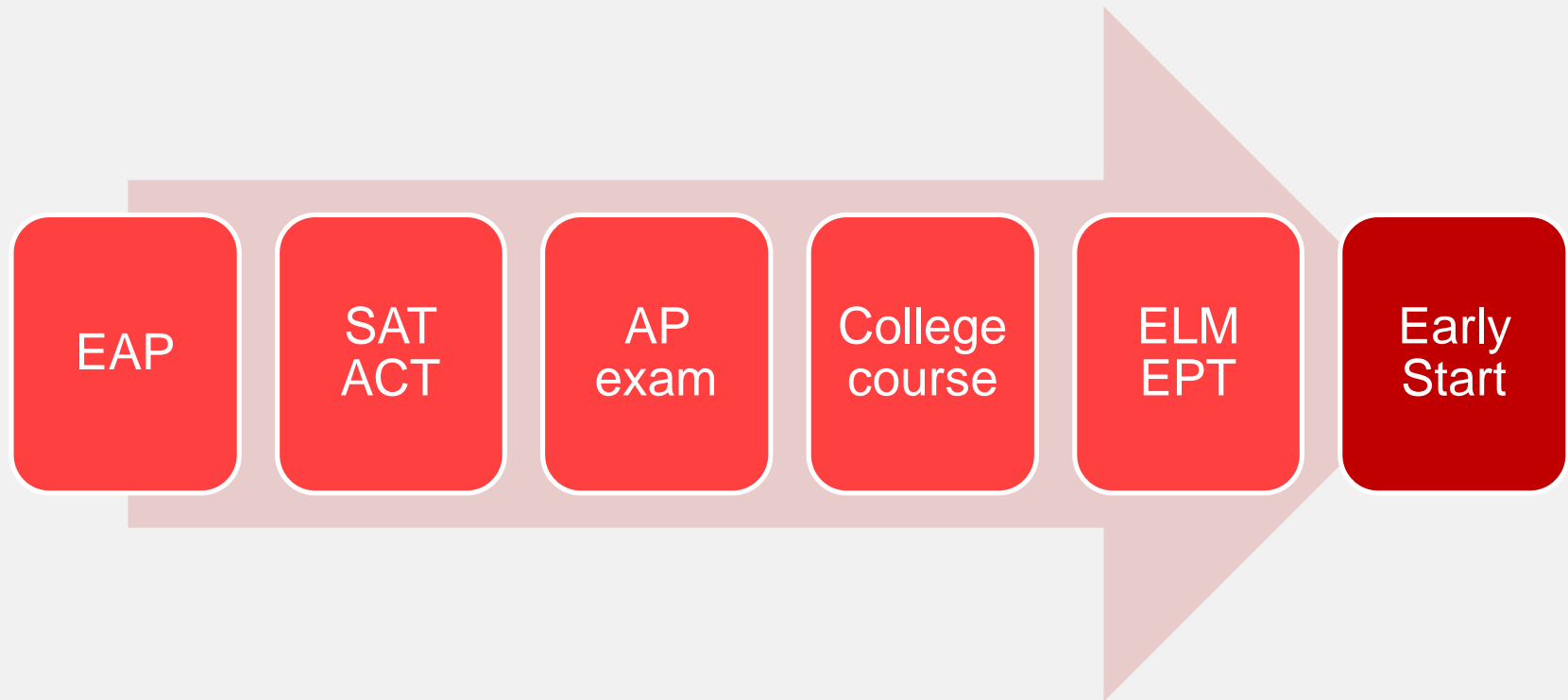
} **4 percentage points**
- Math
  - 80% males
  - 67% females

} **13 percentage points**

# Proficiency: Fall 2014 first-time freshmen



# Demonstrating college readiness at CSU



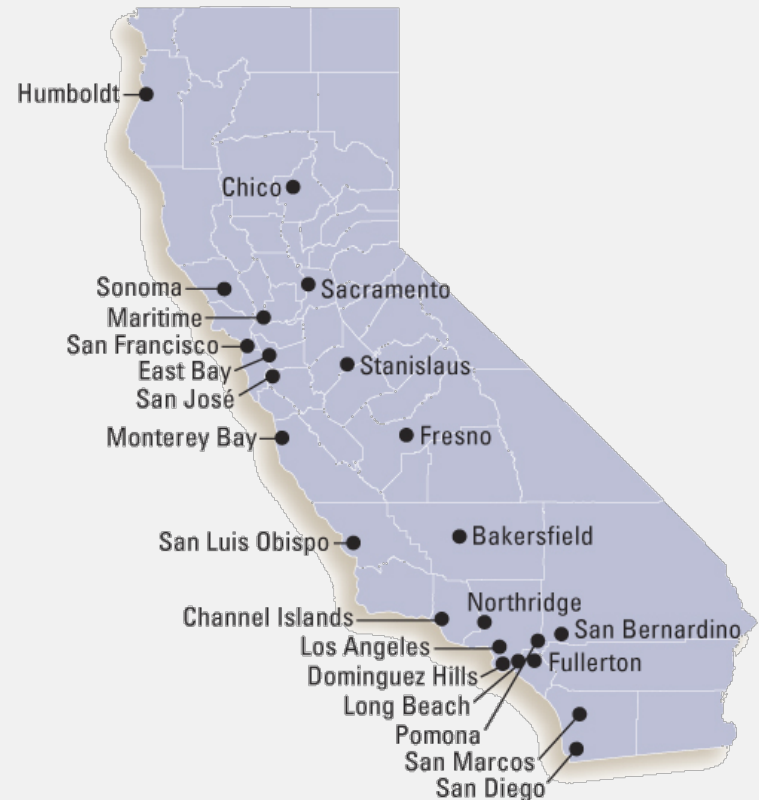


## Early Start Program

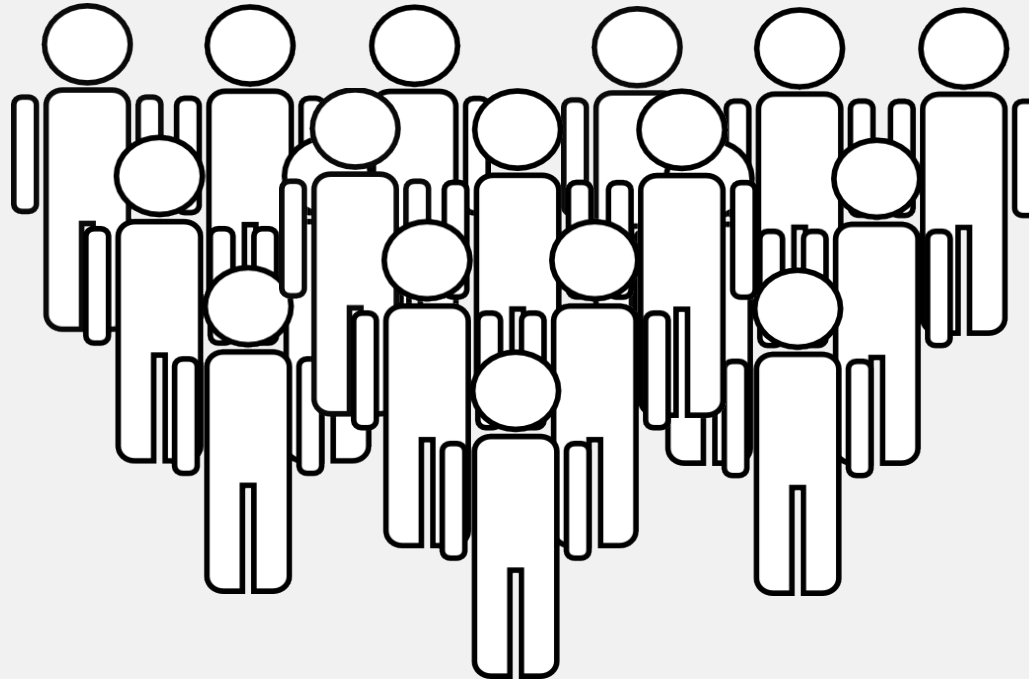
- Required for students who haven't demonstrated college readiness in English and/or math
- Courses summer before freshman year
- Began 2012 (~15,000 students)
- By 2014, mandatory (~24,000 students)
- Goals
  - Reduce pre-college course enrollments
  - Reduce student and institutional cost

## Early Start courses

- Local or destination CSU
- Online or in-person/hybrid
- Credit options
  - 1 unit (intro)
  - 1-2 units (advanced)
  - 3-4 units (most progress)

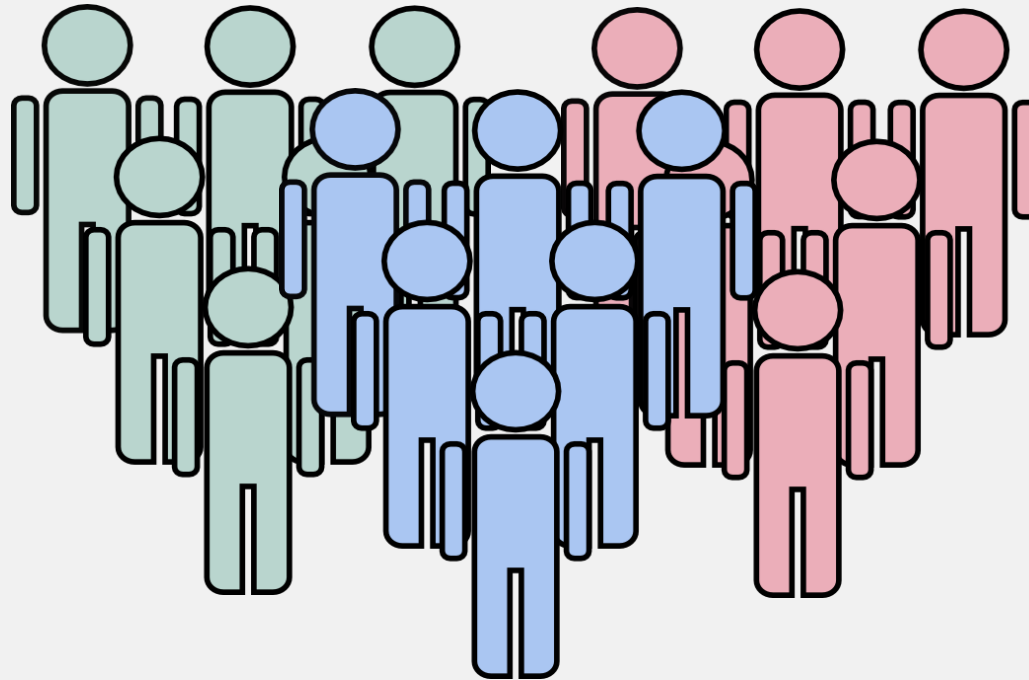


# The issue





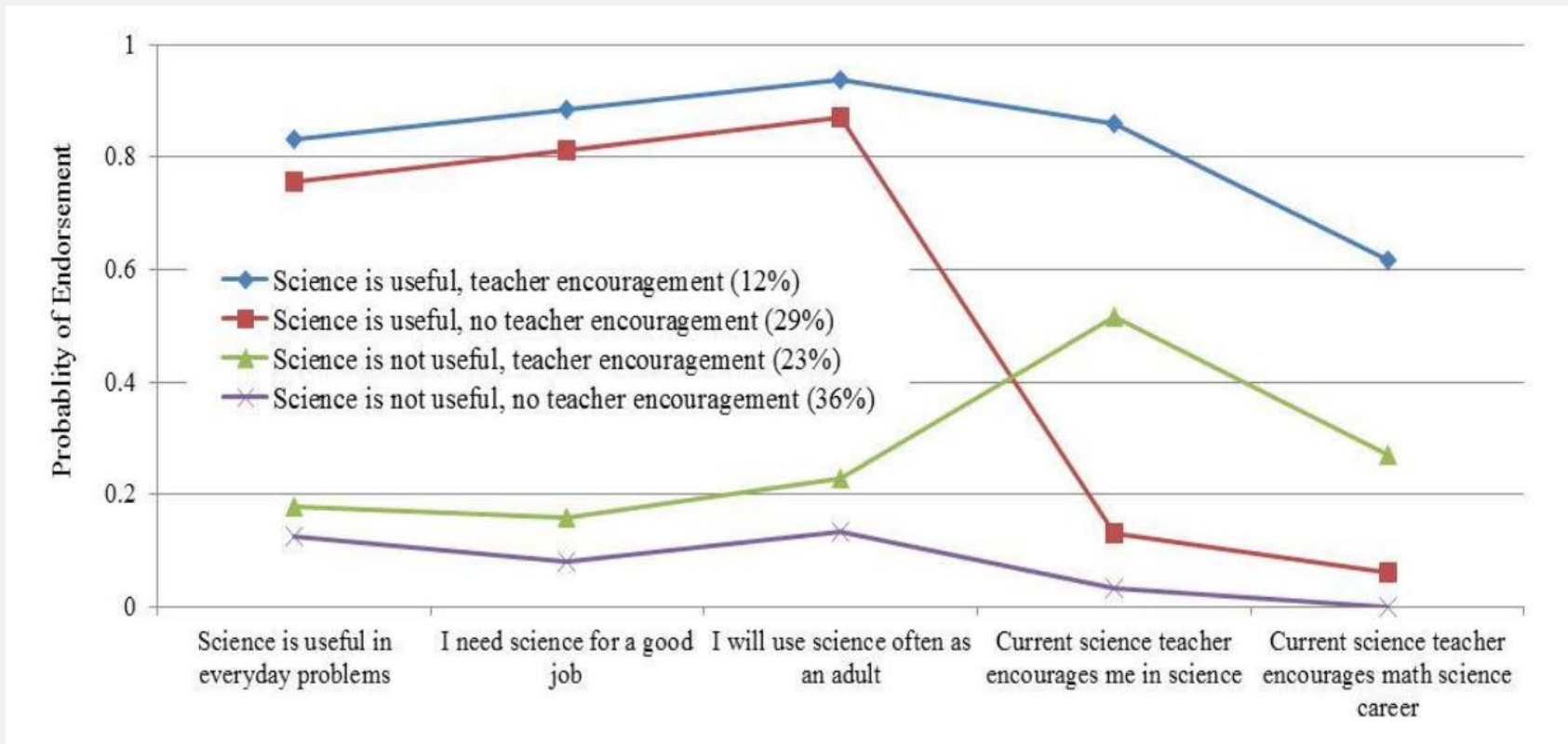
# The issue



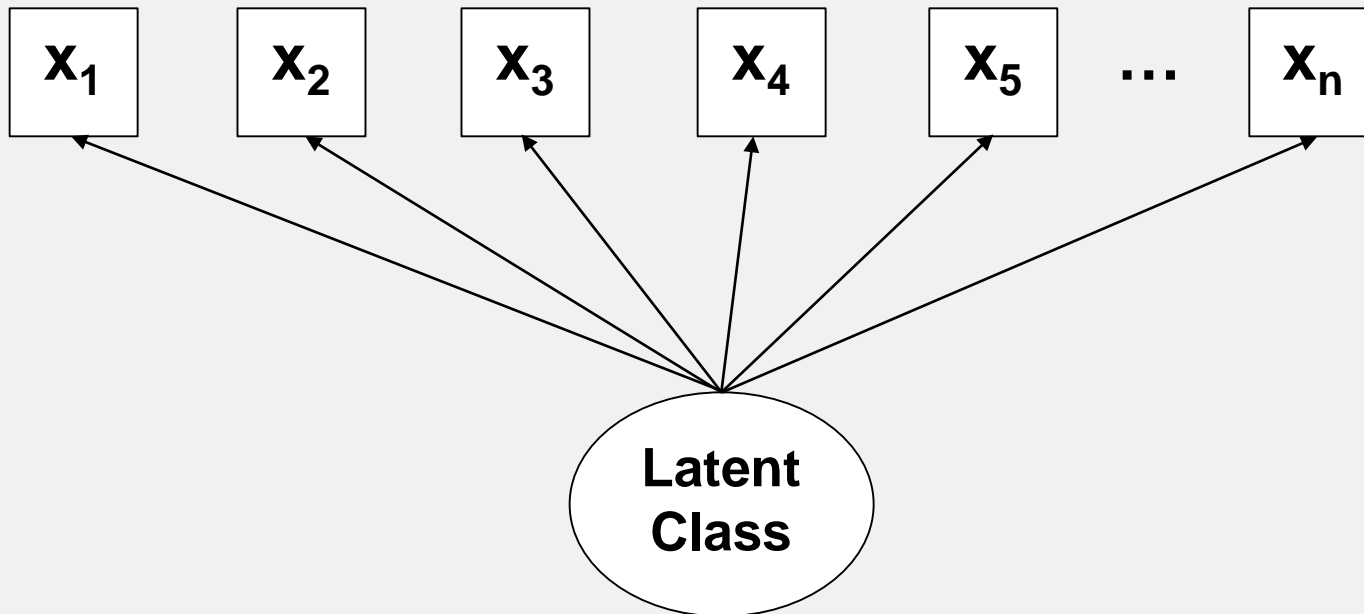
## Latent class analysis (LCA)

- Finds subpopulations within a distribution
- Uses *observable* variables to determine existence of *latent* (unobservable) classes
- Model based: run multiple models, compare
- Based on 2 parameters (max. likelihood)
  - Probabilities of item-response
  - Probability of class membership

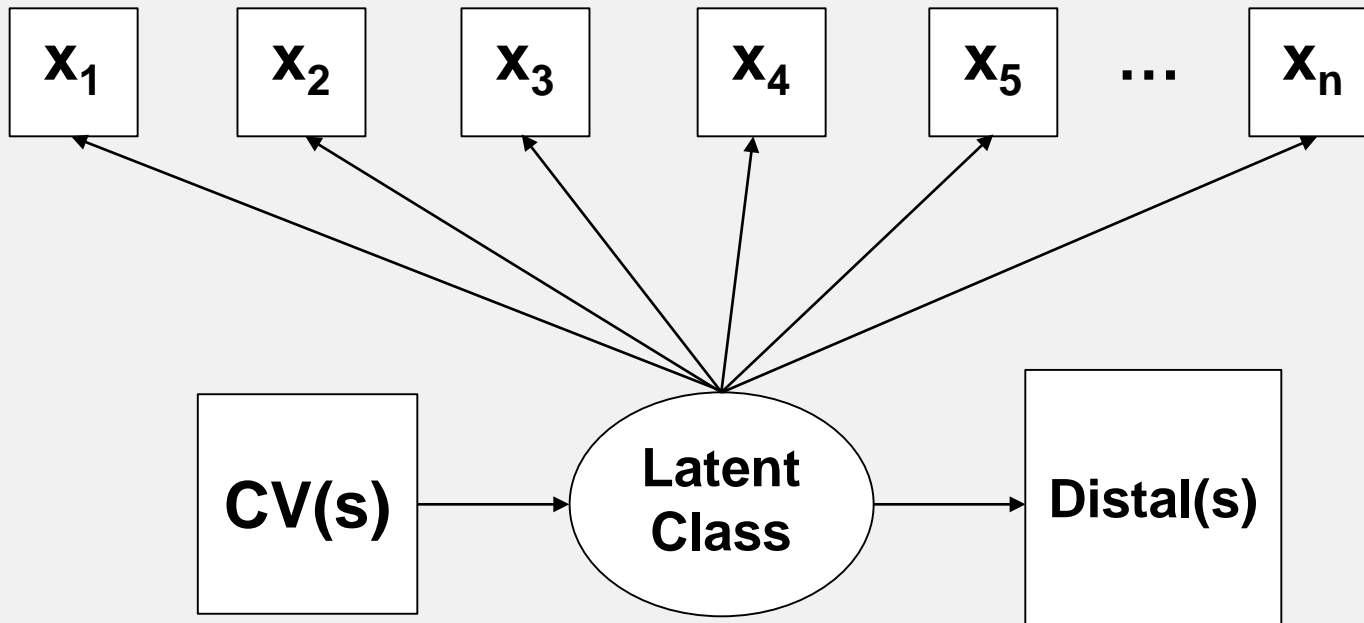
# LCA example: Nylund-Gibson, Ing, & Park (2013)



# LCA model



# LCA model - Covariates, Distal Outcomes



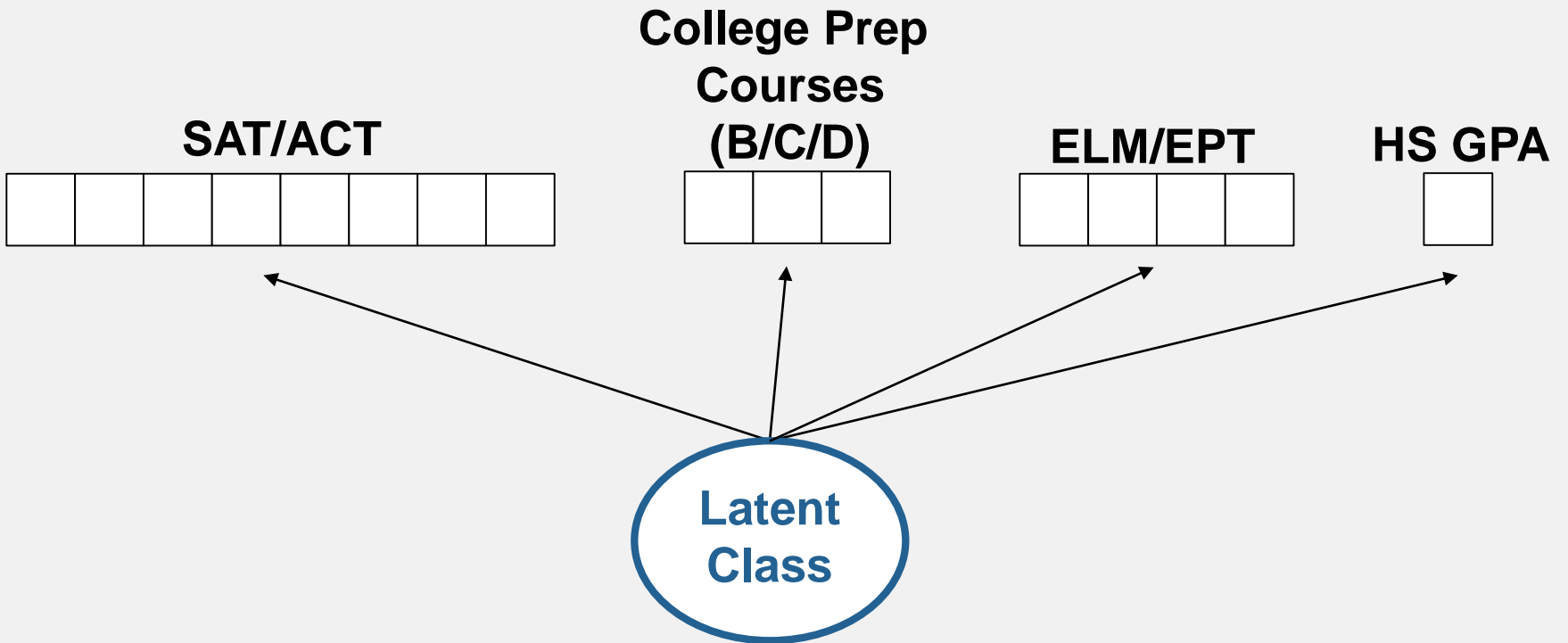
## Very basic LCA how-to

1. Run model for each  $k$  class solution
  - a)  $k = 1$
  - b)  $k = 2$
  - c)  $k = \dots$
2. Compile fit statistics across all models
3. Select best class solution based on fit statistics, theory, and interpretability

## LCA software

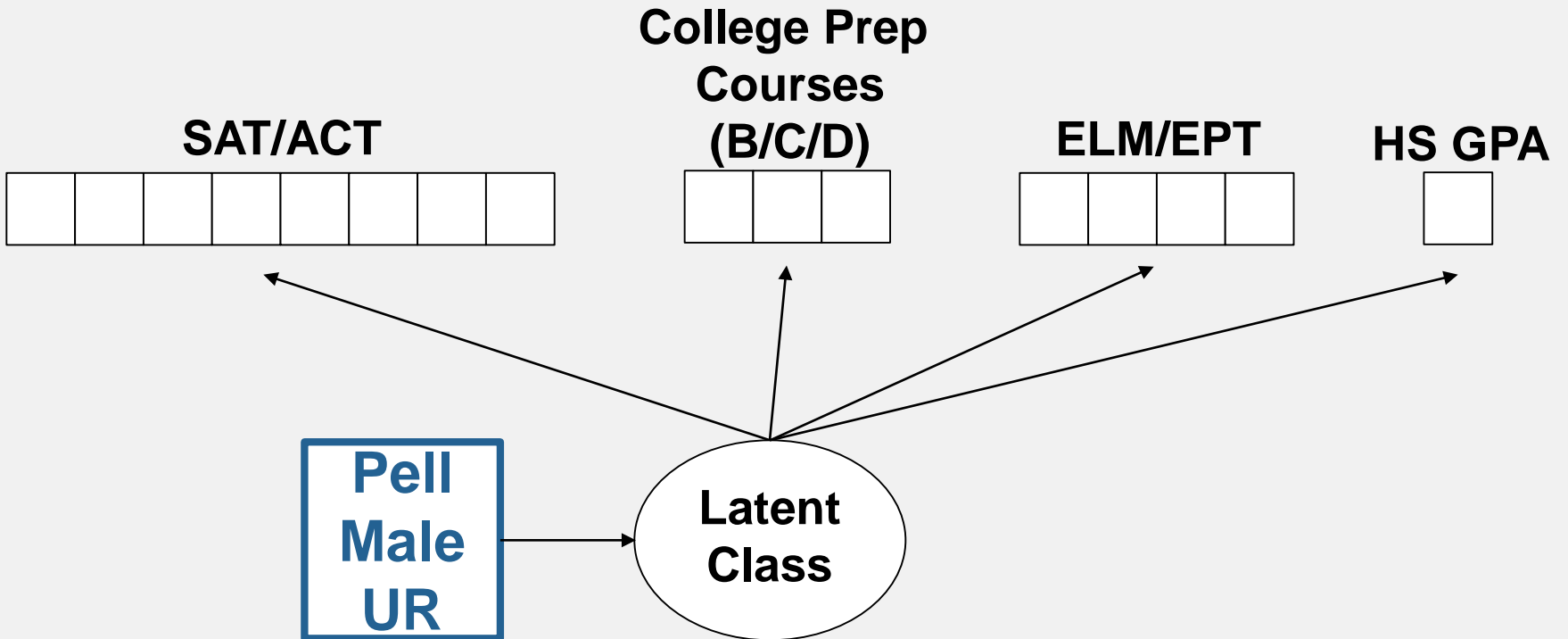


# RQ 1: Are there latent subgroups within ES participants?

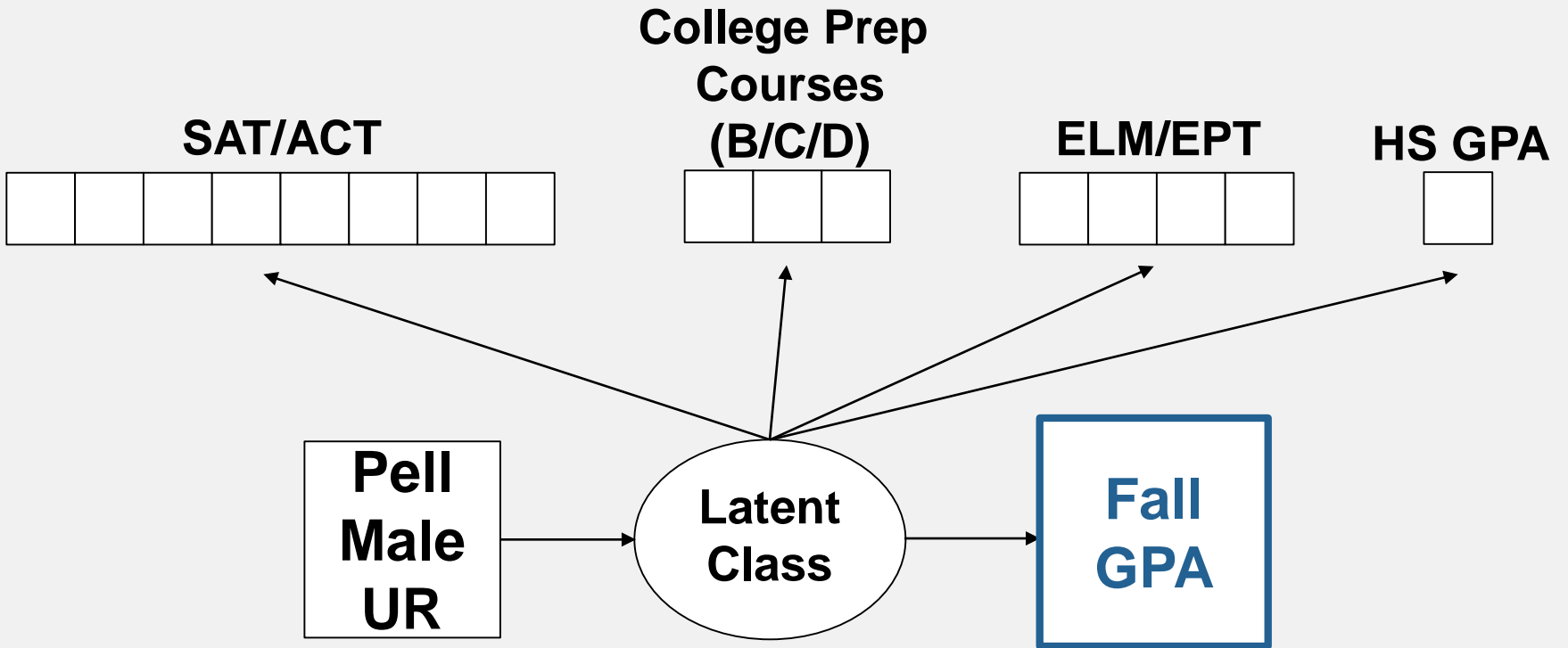




# RQ 2: Who is most likely to be in each class?

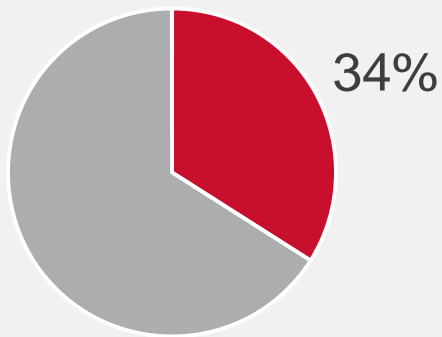


# RQ 3: Which class is most likely to have higher fall GPAs?



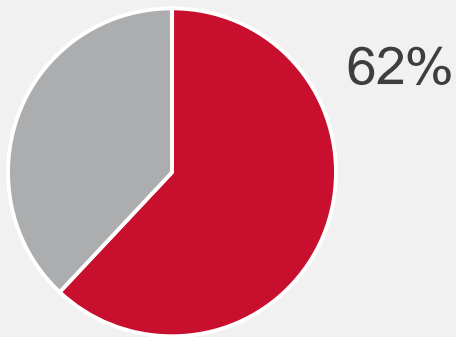
# Summer 2014 Early Start students (*N* = 20,368)

Male



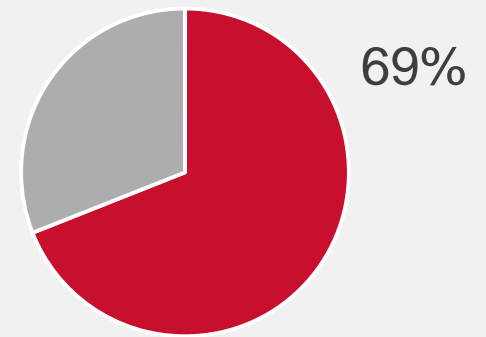
■ Male ■ Female

Pell



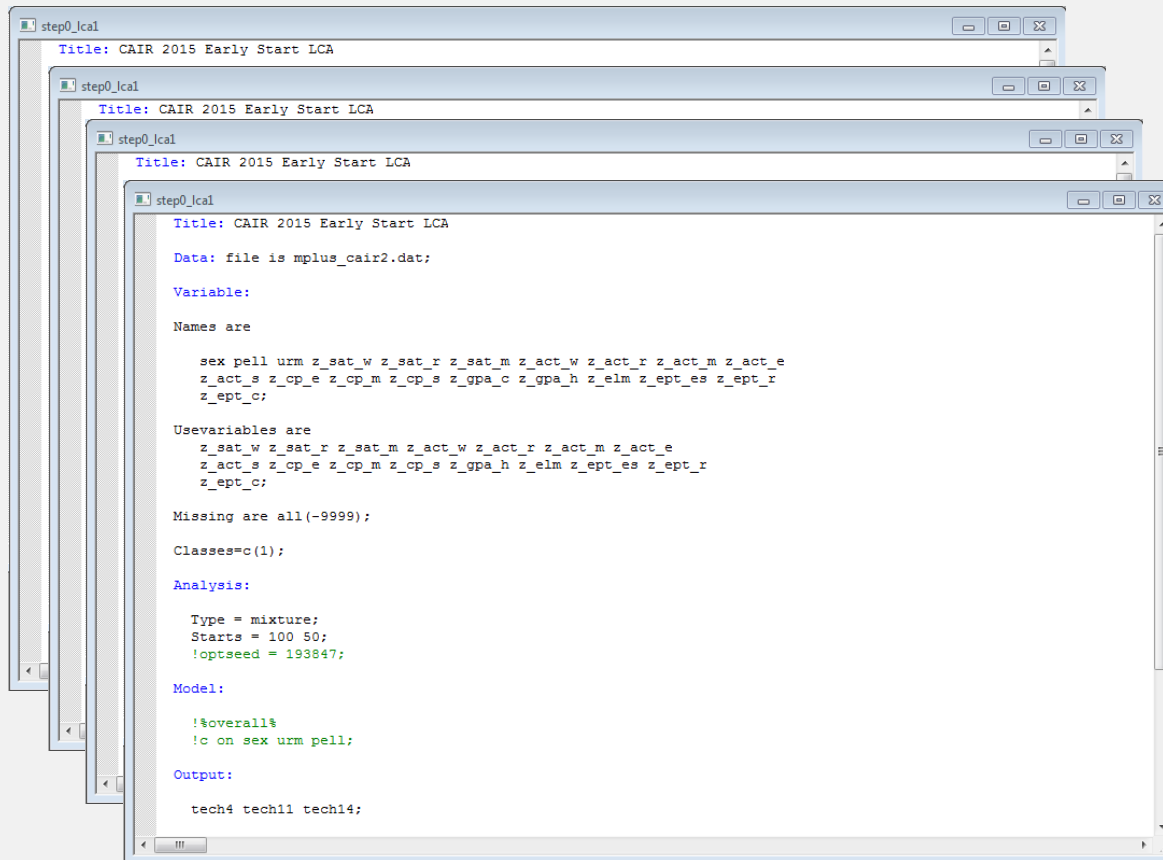
■ Yes ■ No/Unknown

UR



■ UR ■ Non-UR

# Run models



```
step0_lcal
Title: CAIR 2015 Early Start LCA

Data: file is mplus_cair2.dat;

Variable:

Names are

    sex pell urm z_sat_w z_sat_r z_sat_m z_act_w z_act_r z_act_m z_act_e
    z_act_s z_cp_e z_cp_m z_cp_s z_gpa_c z_gpa_h z_elm z_ept_es z_ept_r
    z_ept_c;

Usevariables are
    z_sat_w z_sat_r z_sat_m z_act_w z_act_r z_act_m z_act_e
    z_act_s z_cp_e z_cp_m z_cp_s z_gpa_h z_elm z_ept_es z_ept_r
    z_ept_c;

Missing are all(-9999);

Classes=c(1);

Analysis:

    Type = mixture;
    Starts = 100 50;
    !optseed = 193847;

Model:

    !%overall%
    !c on sex urm pell;

Output:

    tech4 tech11 tech14;
```

# LCA: Selecting the best model

## Comparative fit

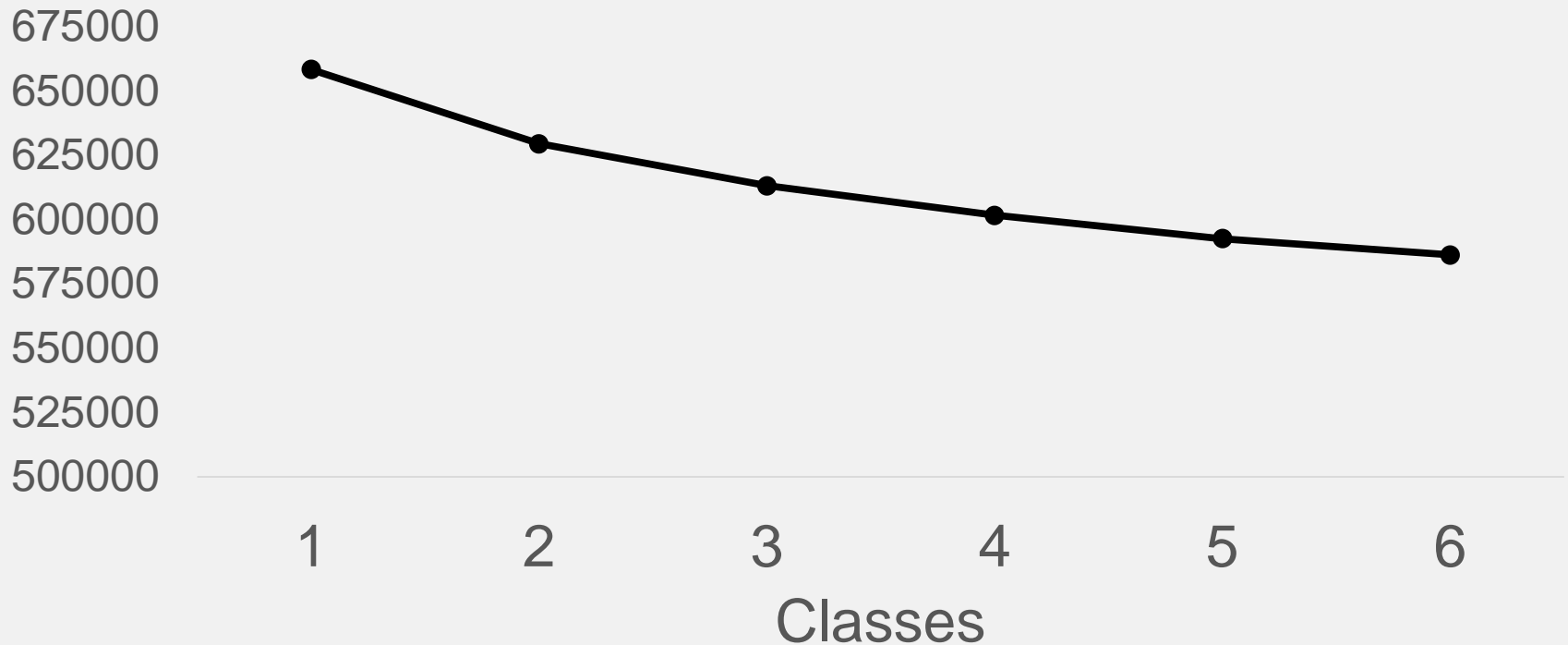
Sig. means  
better than  $k - 1$

Classification  
certainty

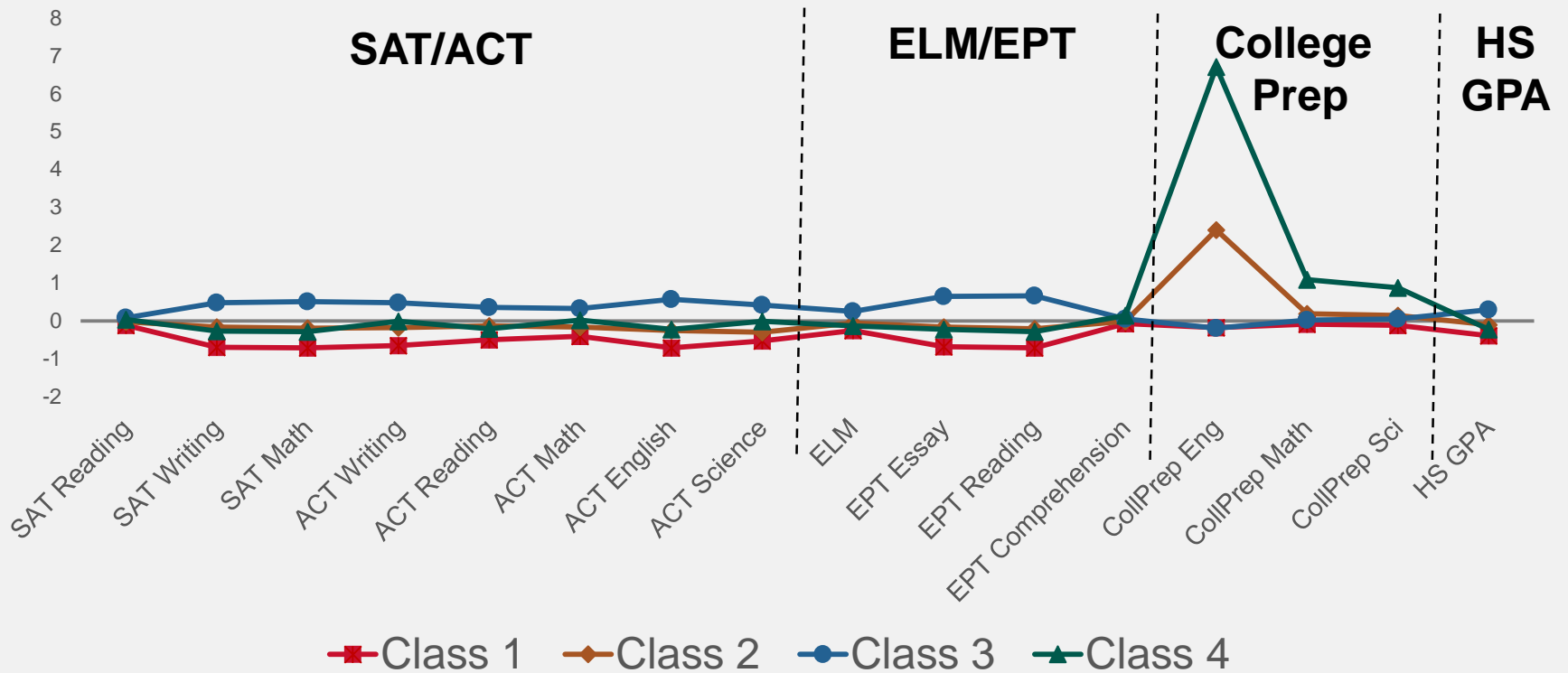
<i>k</i> classes	AIC	BIC	ABIC	BLRT	Entropy
Selection Criterion	Lowest value	Lowest value	Lowest value	Significant value	> .80
<b>1-Class Solution</b>	658018.31	658271.80	658170.11	-	1.00
<b>2-Class Solution</b>	628988.04	629376.21	629220.49	29064.27***	.79
<b>3-Class Solution</b>	612574.21	613097.04	612887.30	16447.83***	<b>.87</b>
<b>4-Class Solution</b>	600995.15	601652.65	601388.88	11613.06***	<b>.87</b>
<b>5-Class Solution</b>	591742.16	592534.34	592216.54	Did not converge	<b>.84</b>
<b>6-Class Solution</b>	<b>585344.66</b>	<b>586271.50</b>	<b>585899.68</b>	Did not converge	<b>.82</b>

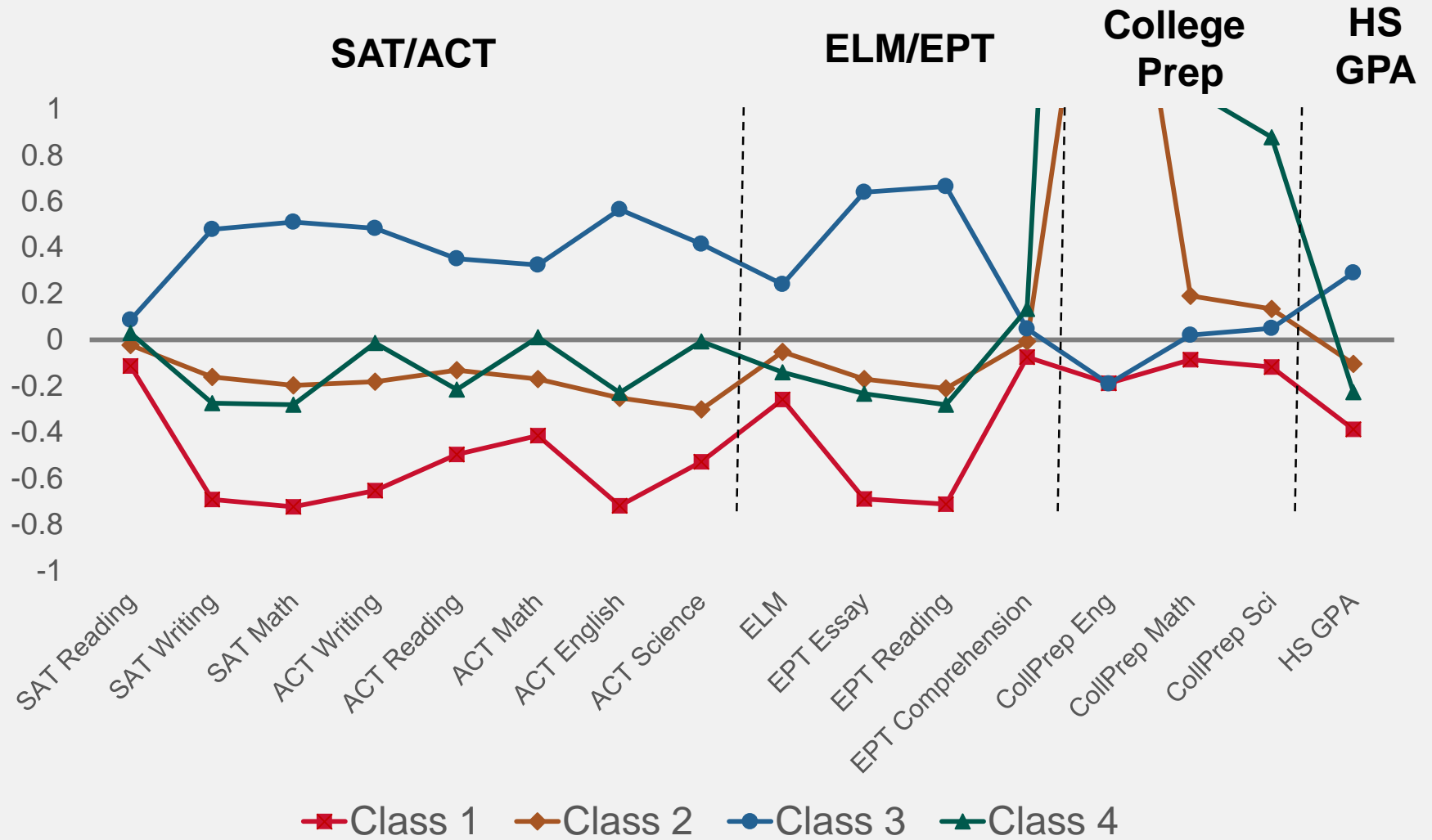
# LCA: Selecting the best model

BIC



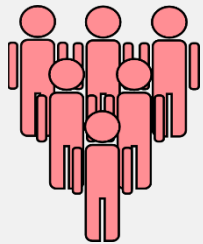
# LCA: Early Start students



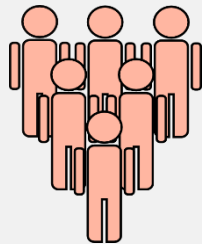




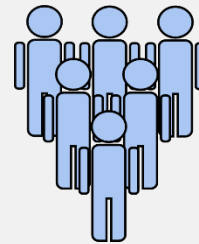
# LCA: Early Start students



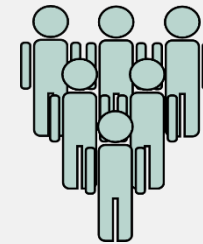
**Class 1 (38%)**  
**Lowest on all**



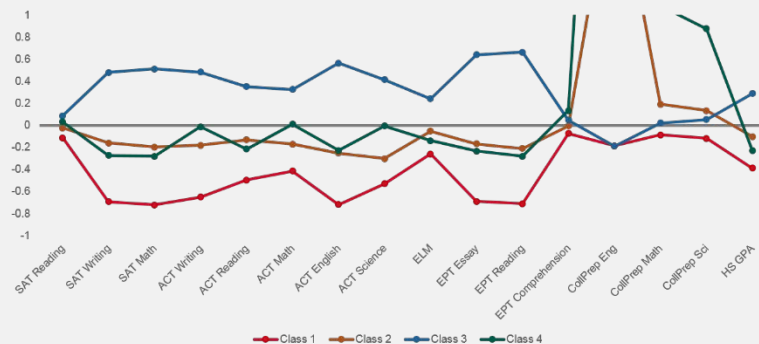
**Class 2 (4%)**  
**Low tests**  
**High B**  
**Low HS GPA**



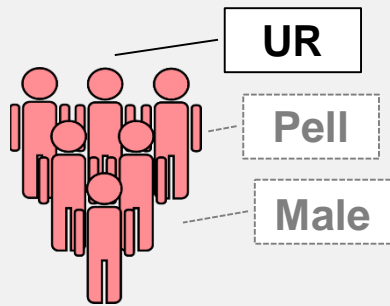
**Class 3 (57%)**  
**High tests**  
**Low B/C/D**  
**High HS GPA**



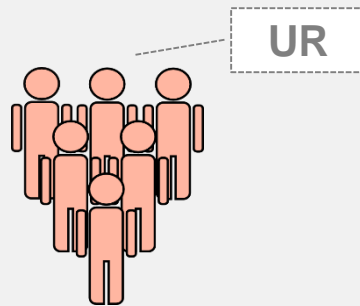
**Class 4 (1%)**  
**Low tests**  
**High B/C/D**  
**Low HS GPA**



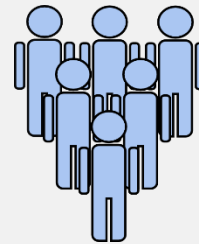
# Covariates



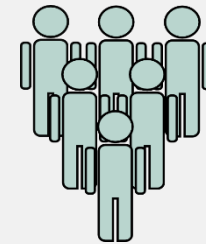
**Class 1 (38%)**  
**Lowest on all**



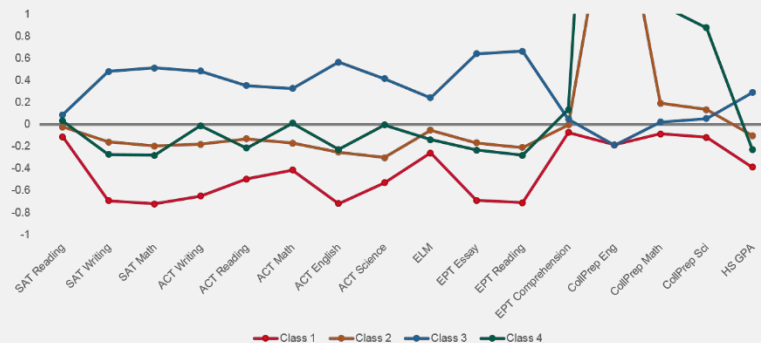
**Class 2 (4%)**  
**Low tests**  
**High B**  
**Low HS GPA**



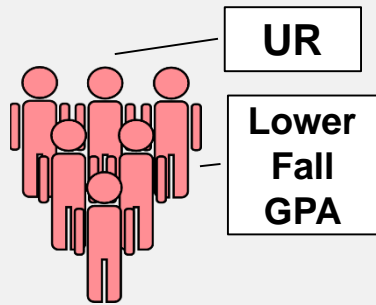
**Class 3 (57%)**  
**High tests**  
**Low B/C/D**  
**High HS GPA**



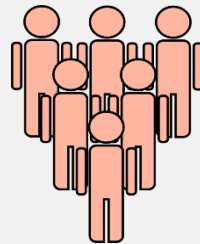
**Class 4 (1%)**  
**Low tests**  
**High B/C/D**  
**Low HS GPA**



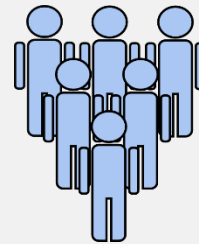
# Outcome: Fall GPA



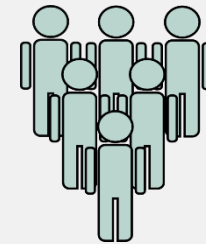
**Class 1 (38%)**  
**Lowest on all**



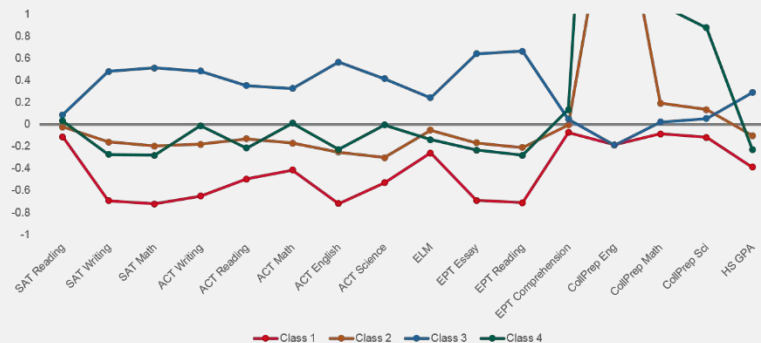
**Class 2 (4%)**  
**Low tests**  
**High B**  
**Low HS GPA**



**Class 3 (57%)**  
**High tests**  
**Low B/C/D**  
**High HS GPA**

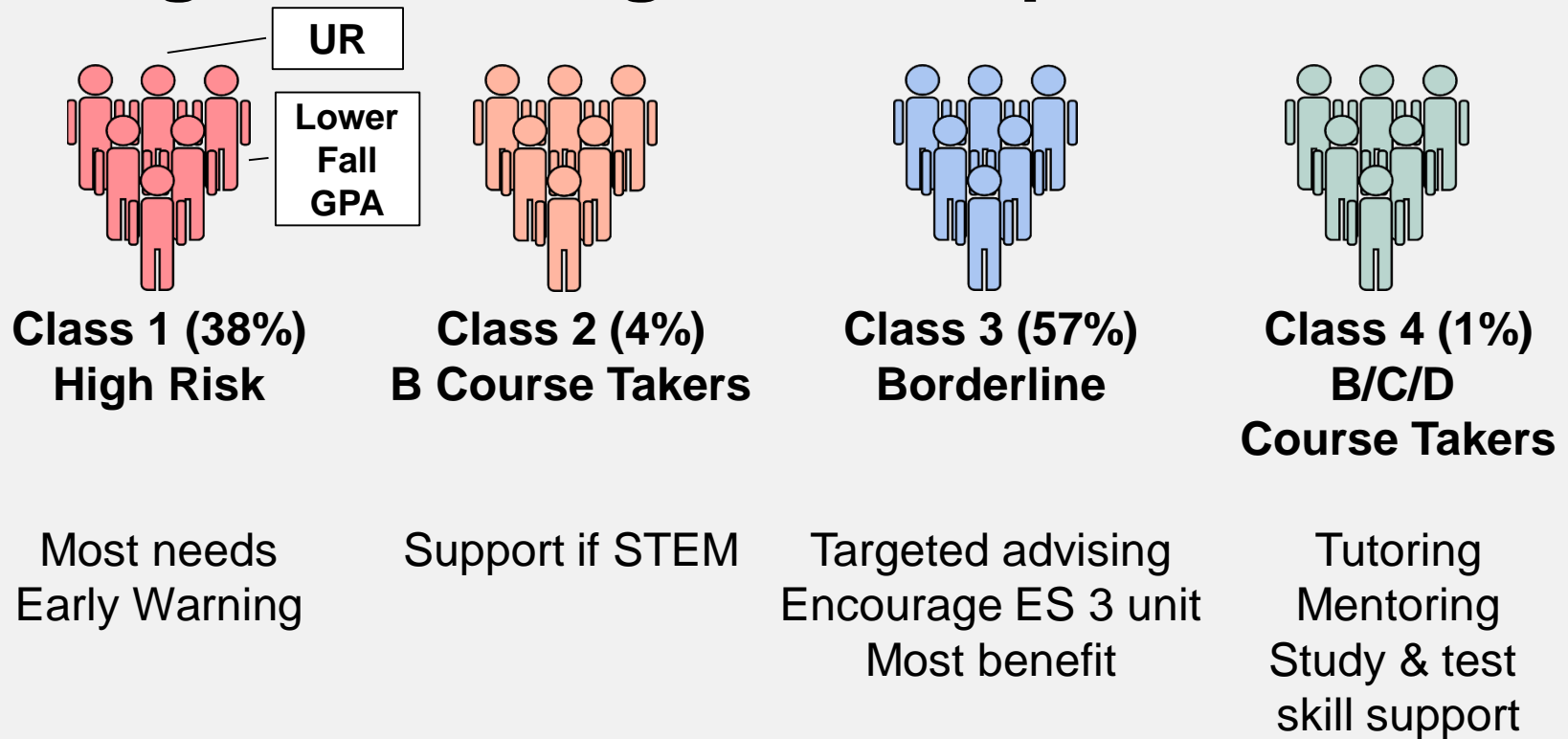


**Class 4 (1%)**  
**Low tests**  
**High B/C/D**  
**Low HS GPA**



	<b>Class 1</b>	<b>Class 2</b>	<b>Class 3</b>	<b>Class 4</b>
SAT Reading	419.3	425.1	<b>432.2</b>	428.6
SAT Writing	388.6	422.1	<b>462.2</b>	414.9
SAT Math	384.7	420.4	<b>468.4</b>	414.7
ACT Writing	6.1	7.6	<b>9.6</b>	8.1
ACT Reading	15.6	17.1	<b>19.2</b>	16.8
ACT Math	16.7	17.5	<b>19.0</b>	18.0
ACT English	13.5	15.3	<b>18.5</b>	15.4
ACT Science	15.9	16.7	<b>19.3</b>	17.8
ELM	35.2	37.3	<b>40.4</b>	36.4
EPT Essay	2.5	2.9	<b>3.5</b>	2.9
EPT Reading	129.1	133.4	<b>141.1</b>	132.8
EPT Comprehension	136.2	136.8	137.3	<b>138.0</b>
College Prep Eng (B)	8.0	10.0	8.0	<b>13.4</b>
College Prep Math (C)	7.6	8.0	7.7	<b>9.3</b>
College Prep Science (D)	6.0	6.4	6.3	<b>7.7</b>
HS GPA	3.1	3.2	<b>3.3</b>	3.1
<b>Fall GPA</b>	<b>2.6</b>	<b>2.8</b>	<b>2.8</b>	<b>2.8</b>

# Using results to guide campus decisions



## Potential LCA studies

- CCC: EOPS entry survey
- UCUES: academic experience, financial situation
- STEM majors
- Veterans
- Online course takers

Erika Baldwin	<a href="mailto:ebaldwin@calstate.edu">ebaldwin@calstate.edu</a>
Matthew Case	<a href="mailto:mcase@calstate.edu">mcase@calstate.edu</a>
Hui-Jeong Woo	<a href="mailto:hwoo@calstate.edu">hwoo@calstate.edu</a>