



Fixing the API

January 18, 2013

Morgan S. Polikoff, Ph.D.
Andrew McEachin, Ph.D.



Overview

- Brief history/background on Academic Performance Index (API)
- **Seven** lessons learned from accountability research
- **Five** steps to improving API
- Q & A



History/background on API

- Public Schools Accountability Act (PSAA, 1999)
 - Created the API
- Dual accountability system due to No Child Left Behind
- Recent signing of SB1458
 - Mandate to revise API
 - Rethinking the role of test scores in measuring school performance



How API is calculated

- Weighted average of students' English language arts (ELA), math, science, and history CSTs
 - Preponderance of weight to ELA and math
- Each student's score on each exam translated to scale
- Scores weighted and averaged to arrive at school API

Performance level	Points
Advanced	1000
Proficient	875
Basic	700
Below Basic	500
Far Below Basic	200



API targets

- State goal is an API of 800
 - Schools above 800 just have to keep it there
 - Schools below 800 must close 5% of the gap between current performance and 800 (e.g., school with API 700 must increase API to 705 next year).
 - Schools also receive subgroup targets calculated in a similar way to the overall targets



API vs. AYP

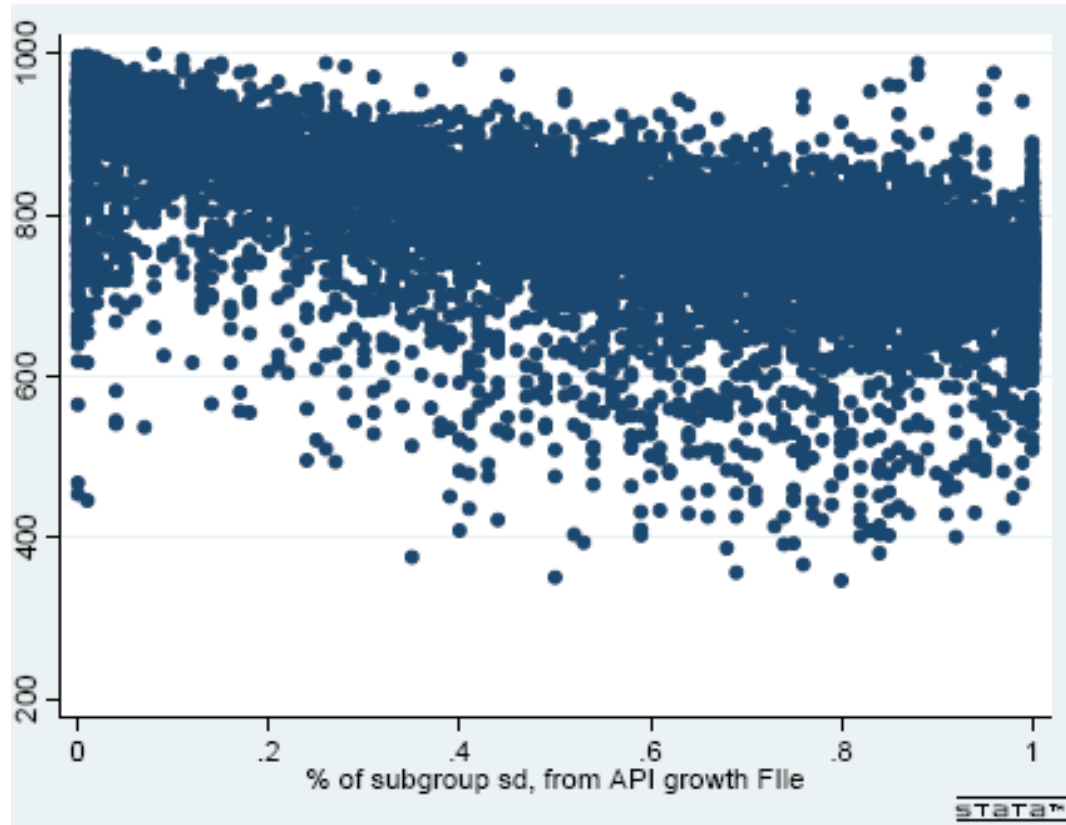
- Under federal No Child Left Behind Act (NCLB), the primary accountability measure is adequate yearly progress
 - Based only on percent of students at or above proficiency
 - Percent of students required to meet proficiency is increasing rapidly to 100% by 2014
 - There are many “loopholes” to avoid these targets, most notably Safe Harbor
 - AYP targets are used to hold Title 1 schools accountable



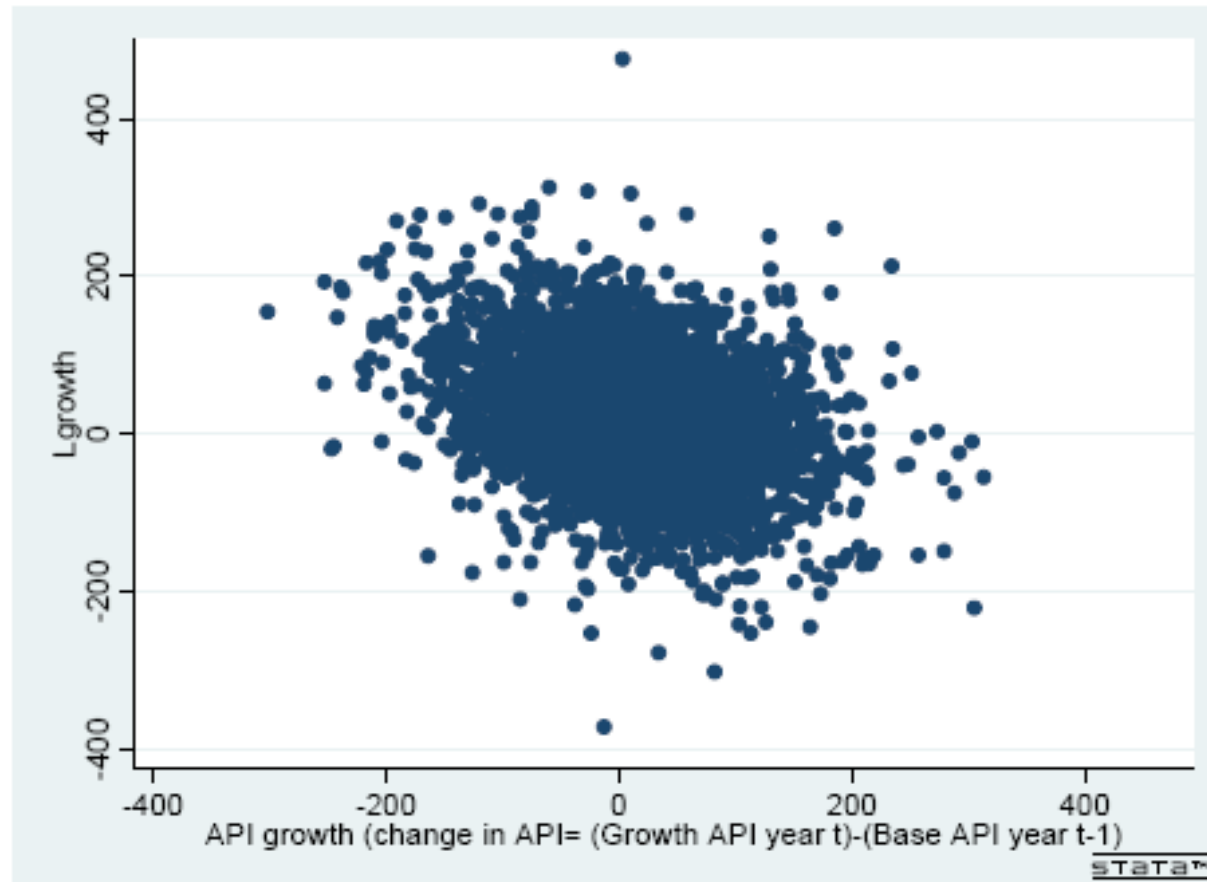
Lessons learned

- API scores primarily measure student demographics
- Change in API from year to year (API growth) is a highly unstable measure
- API and API growth are biased against middle and high schools
- API growth is biased against small schools
- API creates a disincentive to improve achievement for high-achieving students.
- API does not track individual students
- API focuses too narrowly on test scores

API scores largely measure student demographics



Change in API scores is highly unstable

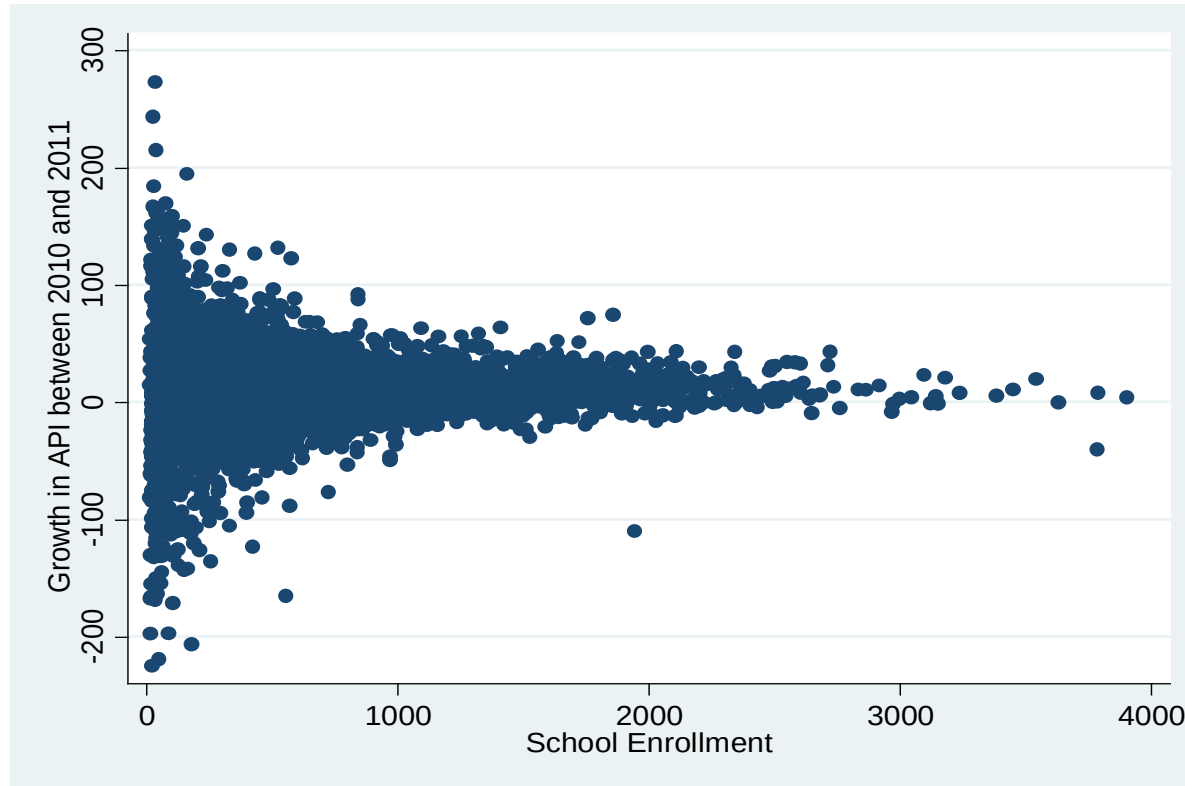


API appears biased against middle/high schools

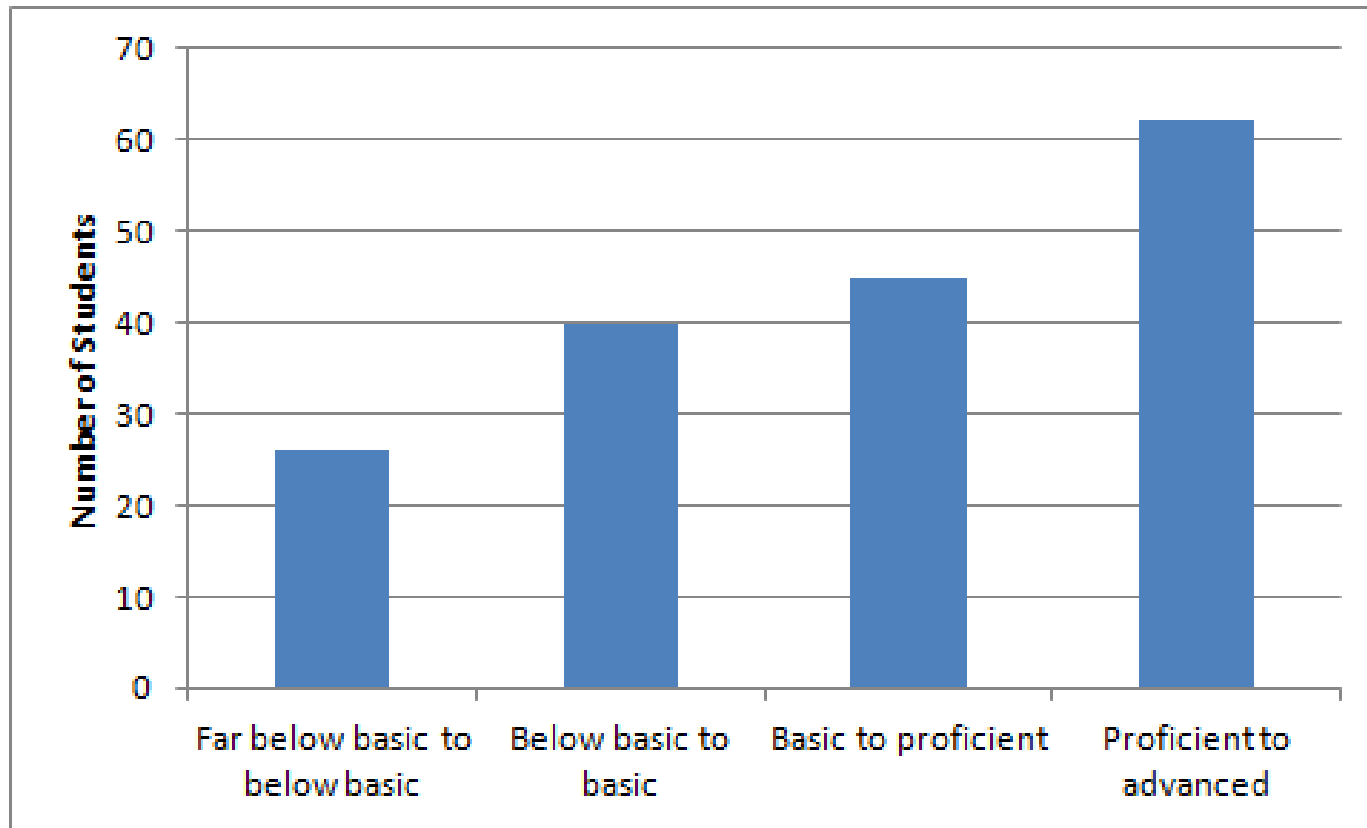


- 45% of elementary schools had API scores below 800 in 2011
 - 58% of middle schools
 - 82% of high schools
- Mean API for elementary schools in 2011 is approximately 810
 - 780 for middle schools
 - 680 for high schools

Growth in API appears biased against small schools



API creates a disincentive to improve achievement for high achievers



API does not track achievement of individual students



	Grade 6	Grade 7	Grade 8	Average
Year one	→ 30	↘ +20		
Year two	→ 30	↘ +20	↘ +20	
Year three	→ 30	↘ +20	↘ +20 → 70	50.0
Year four	→ 30	↘ +20	↘ +20 → 70	50.0
Year five	→ 30	↘ +20	↘ +20 → 70	50.0

API focuses too narrowly on student achievement measured by standardized tests



•While standardized tests are important and predict future life success (see recent piece by Chetty et al.), surely they do not capture everything we want schools to contribute to students. Other potentially important outcomes include:

- Graduation
- Enrollment & persistence in higher education
- Employment
- Students' social well being
- Civic participation / engagement
- and many others ...



Some solutions

- Track the performance of individual students; use these data to replace change in API with a growth/value-added model
- Use multiple years of data to improve the stability of growth data
- Use achievement levels and growth to identify interventions tailored to achievement patterns
- Administer accountability separately by school level and size
- Expand to alternative measures of school performance

Track the performance of individual students



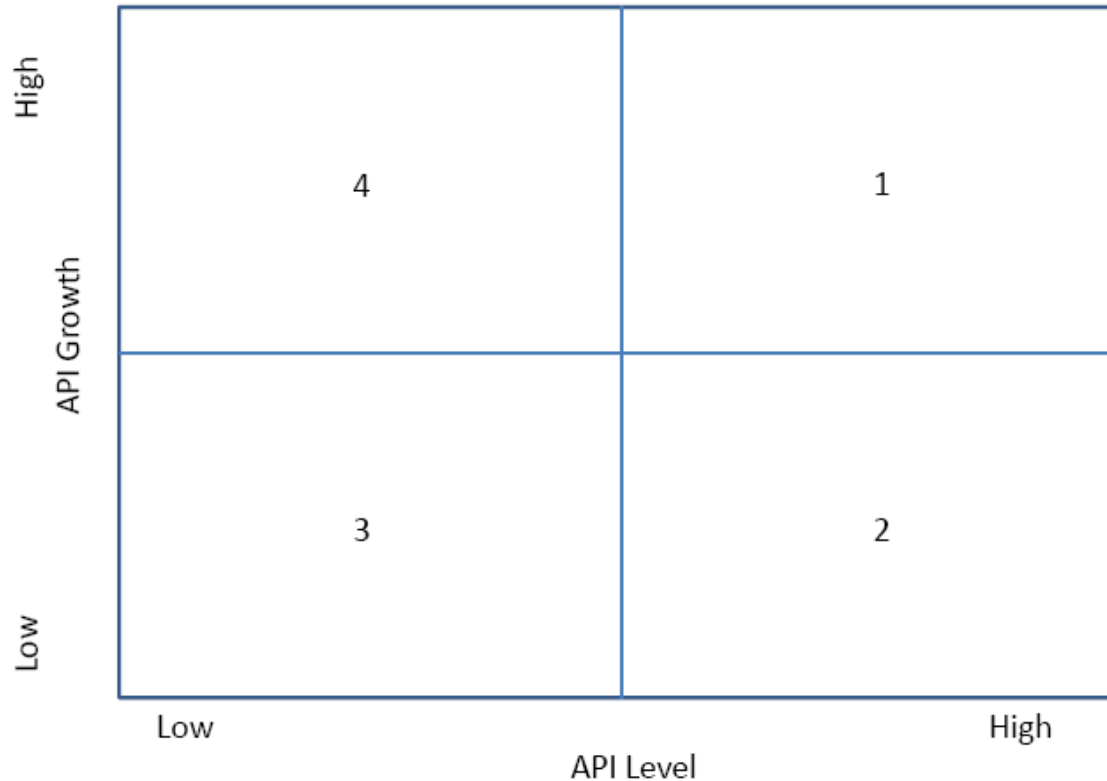
- Smarter Balanced assessments should help with vertical scale, better measurement of high and low achievers
- Longitudinal data systems are already in place
- California need not reinvent the wheel
 - Many states and districts are already using sophisticated methods for analyzing longitudinal student-level data
 - Robust literature on value-added modeling and alternatives
 - VAM experts within the state and outside
 - Need not be used for accountability purposes if that is undesirable

Improve stability of growth measures using multiple years of data



- Reliability of a single year of growth is low
- Combining growth with levels improves reliability
- Averaging growth over a few years – or using data from multiple years without averaging – also improves reliability
- Again, there is a robust literature on which to draw, and experts in the state

Use growth and level to identify interventions tailored to schools



Administer accountability separately by school level and size



- Different school levels will likely have different measures anyway under SB 1458
- Easy to accomplish through stratification
 - For example, select no more than x% of elementary, x% of middle, and x% of high schools for intervention
 - Or, identify the bottom x% by school size, separated into quintiles
 - Can also adjust scores using “shrinkage” techniques to account for instability

Explore alternative measures of school performance



- Robust policy discussion about what are the important outcomes of schooling
- Put our money where our mouth is – measure these important outcomes and include them in our reporting about school performance
- Some of these measures would require new data collection, others would not

Conclusion



- API has served a useful purpose
- The research literature suggests many ways to improve the tool, some of which I have highlighted here
- Any approach should start with identifying desired outcomes of schooling
- Then, measure what's important
- Consider how school performance measures intersect with student and teacher performance measures
- While these techniques will not, on their own, solve our education challenges, a more informative API will be an important step forward



Questions?