

# Teacher Licensure: Theory and Empirical Evidence

Presented at the August 6, 2008

*WA State Basic Education Finance Joint Task Force Meeting*

Dan Goldhaber

Center on Reinventing Public Education

University of Washington

# Overview

- Licensure (certification) system designed to screen out those who don't meet *minimal* quality standards
- System typically entails:
  - Completion of approved teacher training program
    - Pedagogical coursework
    - Student teaching experiences
  - Passing licensure test(s)
- Specific requirements vary considerably by state

# Policy Debates Over Licensure

“It is clear from the evidence that TFA is bad policy and bad education. It is bad for the recruits because they are ill-prepared ... The schools don't get the help they need, and more lasting solutions are not pursued. ... It is bad for the children because they are often poorly taught. ... Finally, TFA is bad for teaching. By clinging to faulty assumptions about what teachers need to know and by producing so many teaching failures, it undermines the profession's efforts to raise standards and create accountability.”

(Linda Darling-Hammond, 1994)

“How many eager, able, nontraditional teacher candidates ... do we lose because of our system of hoops and hurdles and red tape? How many gifted teachers do we lose because they throw up their hands in despair at the obstacles, costs, and coursework between them and the classroom?”

(Finn & Madigan, 2002)

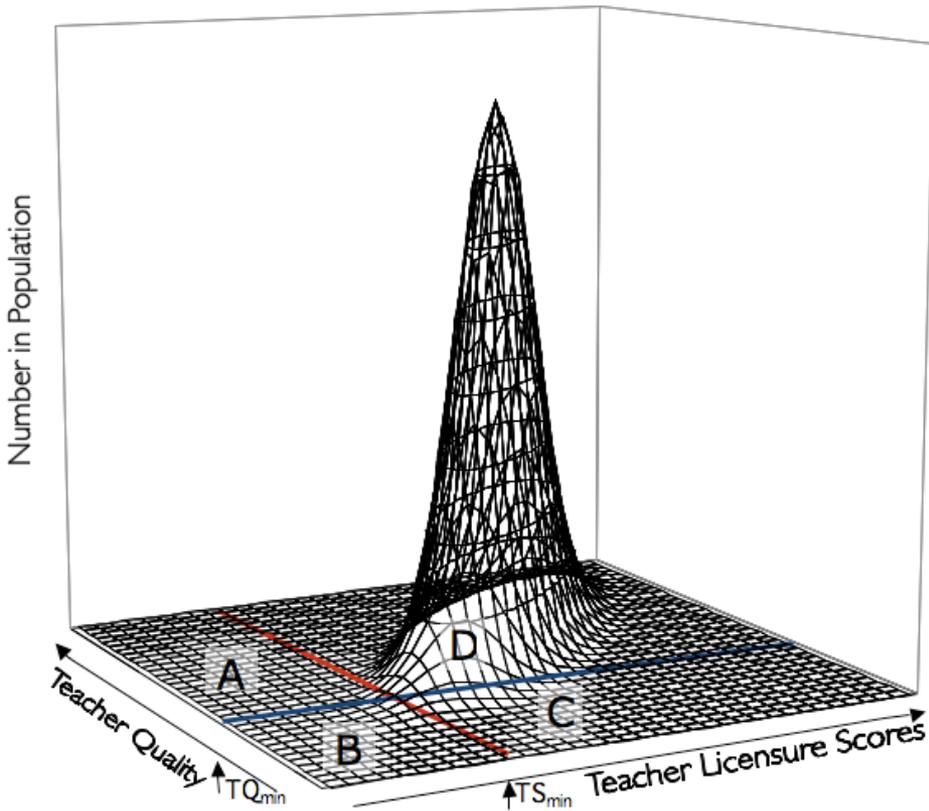
Debate generally framed around “unqualified” versus “unnecessary” barriers: more heat than light

# Impact of Licensure System on Teacher Quality

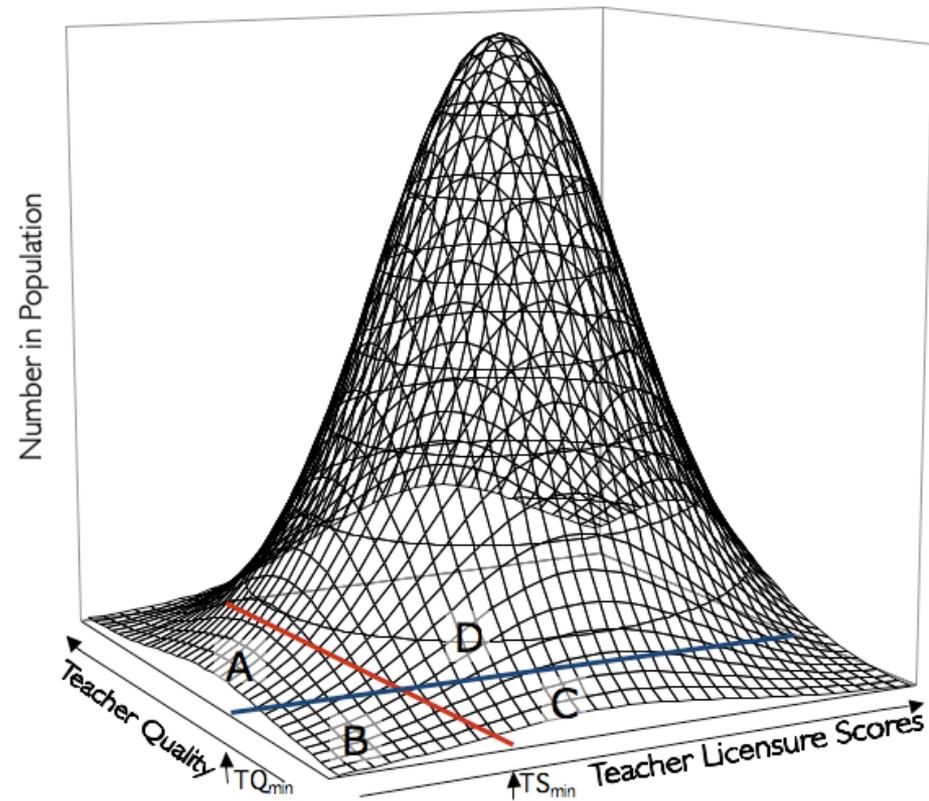
1. How system components affect the *potential* pool of teachers
  - Requirements create costs to individuals (forgone earnings likely to be most important)
2. Strength of the relationship between system (and components such as tests or pedagogical coursework) and student achievement
  - Growing body of empirical literature on this issue
3. Decisions that localities would make with and without state regulation
  - Strongest argument for state regulation is potential for poor local decision-making (nepotism, etc.)

# Relationship Between Components & Student Outcomes Helps Determine Trade-offs

**Panel A**  
Strong Correlation ( $\rho=0.8$ )



**Panel B**  
Weak Correlation ( $\rho=0.2$ )



# Licensure *Tests* & the Potential Pool

- In some states (for example, New Jersey and Texas), alternatively certified teachers represent a substantial proportion of the new teacher workforce
  - TFA received over 20,000 applications for roughly 2,000 spots
- Hanushek & Pace (1995)
  - Analyze teacher training part of pipeline and find that licensure tests reduced number preparing to teach by about 4 percentage points (a reduction of over 30 percent)
- Angrist & Guryan (2008)
  - Analyze teacher workforce composition and find little relationship between teacher tests and workforce composition

*I'm skeptical about both studies*

# Licensure *Status* & Teacher Effectiveness

- Most studies (for example, reviews cited by NCTAF & Abell Foundation) focus on self-efficacy and teaching practices
  - Majority find positive association between these and licensure
- Older literature focuses on student outcomes
  - Hawk, Coble, and Swanson (1985); Rudner (1999); Barnes, Salmon, & Wale (1989), ; Goebel, Romacher, & Sanchez, (1989); Miller, McKenna, & McKenna (1996)
  - Not very credible: not set in value-added framework, no accounting for *relevant alternative*
    - Teachers are not randomly assigned to students so simple comparisons yield unreliable estimates of causal effects
- Only one national licensure status study (Goldhaber & Brewer, 2000)
  - Uses value-added methodology, but relies on self-reports of licensure status and cross-state sample is problematic

# Credible Studies Focused on Impact of Licensed vs. Alternative

- Mathematica random assignment experiment (Glazerman et al., 2005) finds TFA teachers are more effective than other teachers in the *same* school and about as effective as other fully certified teachers in same school
- Boyd et al. (2006) & Kane et al. (2006) studies of various entry pathways (such as TFA, Teaching Fellows, traditional) into NYC schools find little difference in teacher effectiveness based on pathway
  - Average difference in effectiveness (3%) between the most effective (top quintile) and least effective (bottom quintile) teachers *within* a particular category is roughly ten times the average difference in performance *between* teachers who fall into different categories
- Xu et al. (2008) study TFA vs. traditionally licensed teachers and find TFA teachers have large positive effects at the high school level (particularly in science and math)

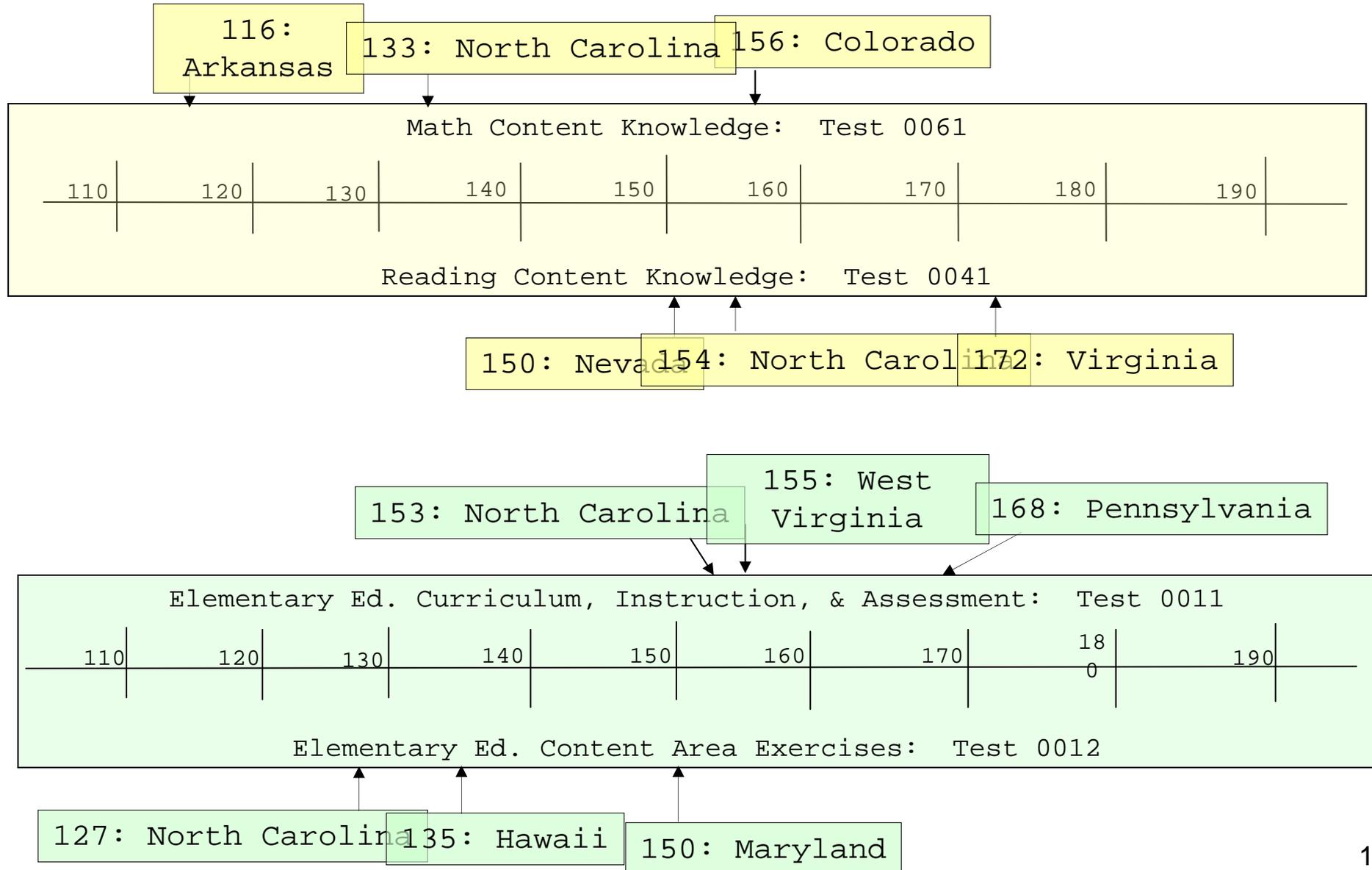
# Licensure Status Overall

- Only a small body of credible evidence on licensure status
  - Generally weak links between licensure and student achievement, but...
  - Licensure status sometimes does appear to predict student achievement (for example, in NC) and sometimes doesn't (for example, in TX); also, the comparison group matters a great deal (not all "alternatives" are equal)
  - Not surprising that licensure would only predict achievement sometimes, since states differ substantially in terms of licensure requirements

# Licensure Components: Credible Quantitative Research

- Teacher training
  - Pre- and in-service pedagogical course requirements
    - Studies rarely find that education degrees predict teacher effectiveness (but, mix of selection and training effects)
    - Jacob & Legren (2004) and Harris & Sass (2007) find mixed results with regard to PD (no positive impact for in-service pedagogical degree)
  - Quality and variation in teacher training institutions (see [www.teacherpolicyresearch.org](http://www.teacherpolicyresearch.org)): linking training components to student achievement
- Teacher licensure testing
  - Wide variety of tests (subject, form, level)
  - Disparate impact of teacher testing - minority teacher candidates tend to perform significantly less well on licensure tests, so higher cutoffs have implications for the diversity of the teacher workforce

# Significant Variation in State Cutoff Scores



# Generally Positive Evidence on Licensure Test Performance & Student Achievement

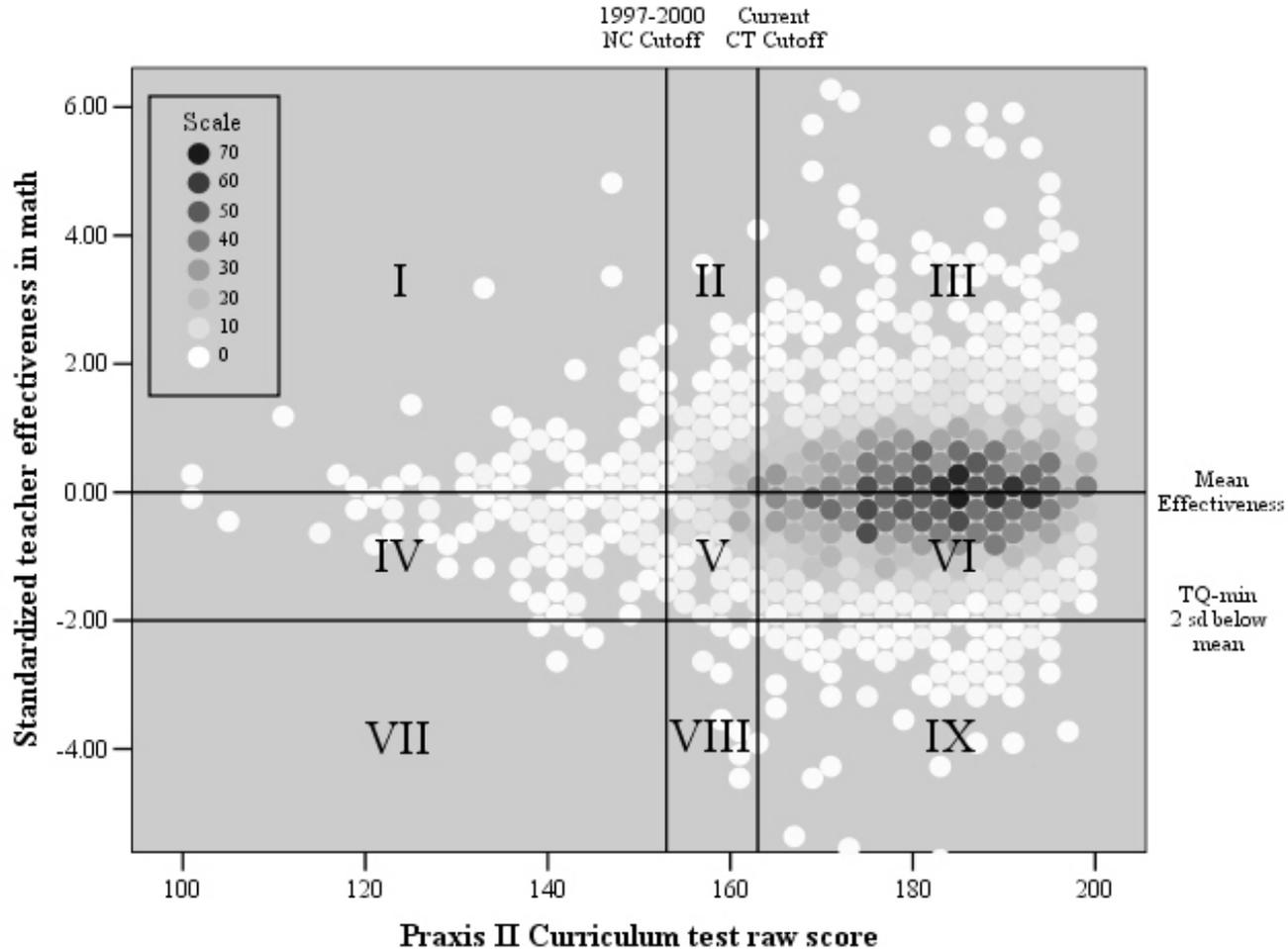
- **Summers & Wolfe (1977)**: Small negative relationship between NTE scores and increase in elementary school test scores
- **Strauss & Sawyer (1986)\***: 1 percent increase in teacher quality results in 5 percent decline in student failure rate
- **Ferguson (1991)\***: 1 s.d. increase in teacher licensure scores raises student scores by 0.17 s.d.
- **Ferguson & Ladd (1996)\***: 1 s.d. change in teacher test scores lead to 0.10 s.d. difference in student test scores
- **Clotfelter, Ladd, & Vigdor (2004, 2007)**: 1 s.d. increase in teacher licensure scores increases predicted student achievement by 1 to 2 percent of a standard deviation
  - Accounts for non-random sorting of students across schools, but not within schools

\* *May suffer from aggregation bias*

# Goldhaber (2007) on Licensure Tests

- Accounts for nonrandom match between teachers and students and exploits changes in state (NC) standards
- Robust findings show that some licensure tests are predictive of teacher effectiveness, but ...
  - Magnitude of estimated coefficients are greatly influenced by teacher sorting across schools and classrooms
- Licensure tests have both screening and signal value
  - Performance across the licensure test distribution (not just at cut-point) predicts teacher effectiveness, but (anecdotally) school systems typically do not use licensure scores when making hiring decisions
  - Increase in cutoff to CT standard does not increase predictive power of pass/fail screen
- Licensure testing is not without costs - there are a large number of false negatives and positives

# Observed Relationship Between Teacher Licensure-Test Performance & Estimated Teacher Effectiveness

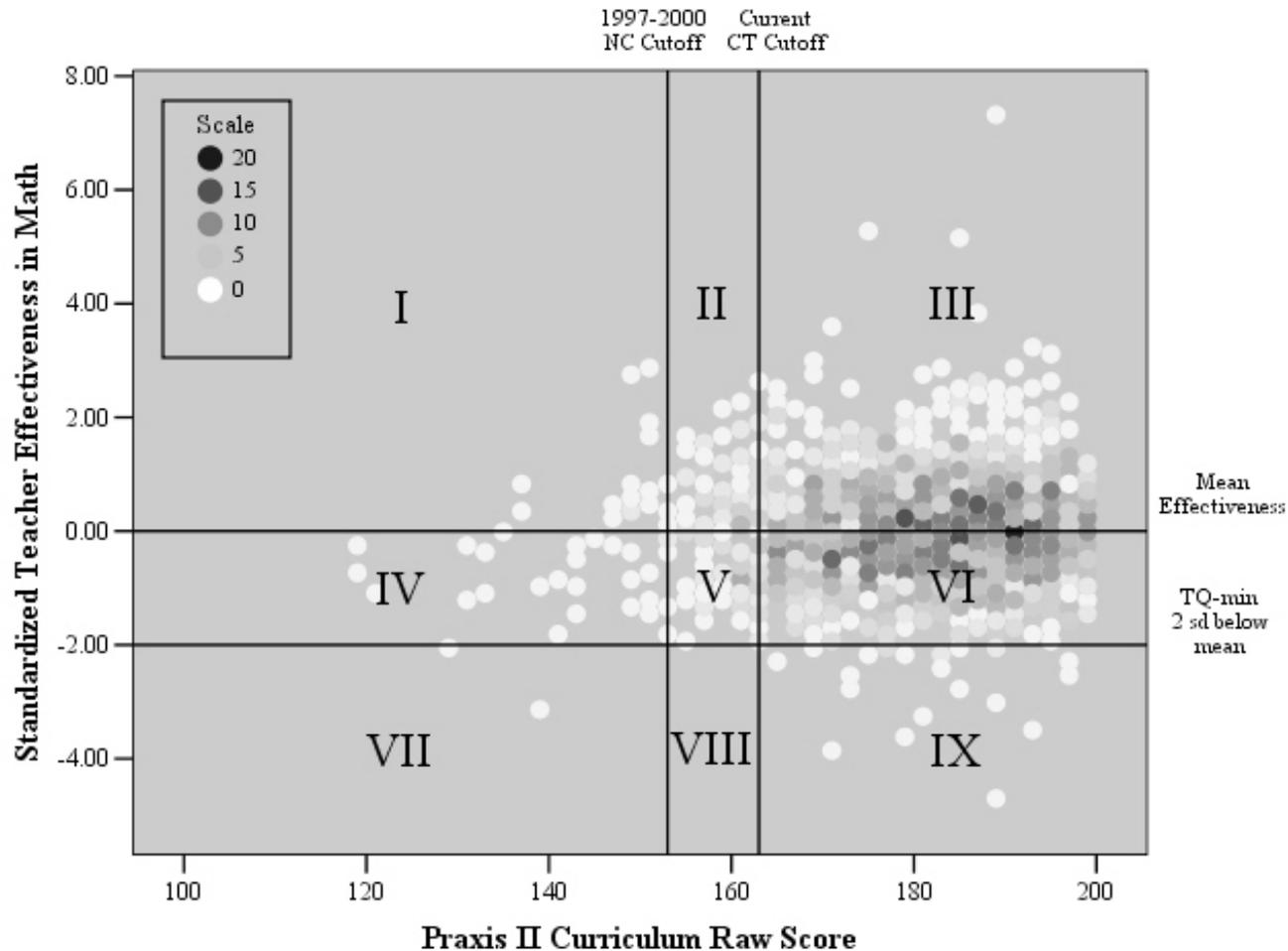


Percentage of sample in each section:

I:	1.2%	II:	3.6%	III:	44.1%
IV:	2.1%	V:	3.6%	VI:	43.5%
VII:	0.1%	VIII:	0.2%	IX:	1.6%

Note: Darker Shades indicate greater density of observations.

# Observed Relationship Between Teacher Licensure-Test Performance & Teacher Effectiveness: 3-year gains



Percentage of sample in each section:

I:	0.9%	II:	3.7%	III:	44.2%
IV:	1.5%	V:	3.3%	VI:	44.9%
VII:	0.1%	VIII:	0.0%	IX:	1.2%

Note: Darker Shades indicate greater density of observations.

# Final Thoughts

- Efficacy of licensure system depends on:
  - Who is screened out by licensure hurdles (is there really a “reserve army” of teachers?)
  - One’s view of local capacity to make judgments about teacher candidates
  - Labor market conditions
  - Value judgments about false negatives vs. harm done by potentially poor teachers
- Likely distributional consequences of licensure system:
  - Within-state variation in applicant pools/hiring capacities
  - In the absence of other changes (such as salary), more rigorous requirements will restrict supply, thus exacerbating differences in applicant pools

# Backup Slides

# Goldhaber & Brewer Research

- Data
  - *National Educational Longitudinal Survey of 1988* (NELS:88)
    - Link students directly to individual teachers and classes
  - State certification policies 1990 (for example, cutoff scores)
- 12th grade science and math test scores regressed on:
  - Individual and family background characteristics
  - School, teacher, and class level variables
  - Teacher certification status
    - Standard, probationary, emergency, private school, not certified in subject

# Sample Statistics: Student Variables

	Standard Certification in Subject		Probationary Certification in Subject		Emergency Certification in Subject		Private School Certification		Not Certification in Subject	
	Math	Sci	Math	Sci	Math	Sci	Math	Sci	Math	Sci
12 <sup>th</sup> grade test score	51.52 (13.49)	25.26 (5.87)	41.93 (12.88)	23.50 (6.67)	43.74 (13.76)	23.29 (5.12)	50.01 (14.35)	25.07 (6.87)	41.94 (13.03)	24.20 (5.02)
10 <sup>th</sup> grade test score	46.47 (13.01)	23.39 (5.76)	35.42 (11.51)	22.38 (5.46)	37.94 (12.44)	20.61 (5.01)	46.42 (14.29)	23.80 (6.55)	38.35 (12.84)	22.46 (5.57)
Gain 10 <sup>th</sup> to 12 <sup>th</sup>	5.05 (5.36)	1.87 (3.71)	6.51 (3.54)	1.12 (3.53)	5.80 (5.12)	2.68 (3.58)	3.59 (5.95)	1.26 (3.95)	3.59 (4.68)	1.74 (3.47)
M other's education	13.00 (2.92)	13.16 (2.94)	12.28 (2.99)	13.03 (2.98)	12.33 (2.92)	12.59 (3.26)	13.43 (2.83)	12.69 (3.78)	11.83 (3.04)	13.00 (2.65)
Father's education	13.41 (3.24)	13.77 (3.31)	12.19 (3.62)	12.58 (3.39)	11.19 (3.11)	12.82 (4.21)	12.83 (3.24)	14.40 (3.57)	12.53 (2.93)	13.16 (2.91)
M other not in househld	.06	.04	.13	.02	.10	.05	.12	.11	.09	.14
Family income	46681 (38133)	50000 (38472)	35785 (19527)	42689 (30745)	34168 (23575)	48117 (54838)	50875 (38358)	48986 (45346)	31453 (22762)	38269 (20106)
<b>Number of students</b>	3179	2069	24	41	49	41	58	44	77	29

# Sample Statistics: Teacher Variables

	Standard Certification in Subject		Probationary Certification in Subject		Emergency Certification in Subject		Private School Certification		Not Certification in Subject	
	Math	Sci	Math	Sci	Math	Sci	Math	Sci	Math	Sci
% minority students in class	27.82 (31.30)	25.68 (29.61)	44.59 (39.60)	32.26 (37.33)	33.00 (36.75)	35.25 (31.79)	22.22 (19.32)	24.68 (34.34)	42.42 (35.73)	27.50 (26.75)
Teacher white	.91	.93	.71	.82	.88	.88	.97	.88	.75	.89
Teacher has MA	.58	.55	.20	.30	.21	.37	.61	.48	.52	.32
BA major in subject	.75	.49	.68	.83	.57	.53	.74	.83	.13	.11
BA major in education	.29	.37	.16	.09	.24	.32	.39	.22	.49	.50
MA major in subject	.43	.49	0	.29	.57	.43	.48	.62	.12	.13
MA major in education	.45	.37	.60	.71	.29	.29	.43	.15	.48	.50
<b>Number of teachers</b>	1695	1106	21	23	34	24	34	27	52	21

# Summary of Descriptive Statistics

- 10th- to 12th-grade gain is not higher for teachers with standard certification
- Students with non-standard credentialed teachers:
  - Have lower 10th grade scores
  - Have lower family income & parental education
- Teachers with non-standard credentials:
  - Are more likely to be minority
  - Have less experience
  - Mixed picture on degrees

# Results

	Basic Models		Interaction Models	
	Mathematics	Science	Mathematics	Science
BA major in subject	<b>.41 (1.58)</b>	<b>-.16 (.70)</b>	<b>.42 (1.60)</b>	<b>-.14 (.61)</b>
BA major in education	<b>-.44 (1.76)</b>	<b>-.02 (.11)</b>	<b>-.42 (1.73)</b>	<b>-.09 (.44)</b>
MA major in subject	<b>.58 (1.81)</b>	<b>.23 (.81)</b>	<b>.57 (1.77)</b>	<b>.23 (.90)</b>
MA major in education	<b>-.09 (.29)</b>	<b>-.10 (.34)</b>	<b>-.10 (.31)</b>	<b>-.14 (.55)</b>
Probationary certification in subject	<b>1.29 (1.14)</b>	<b>-.52 (.80)</b>	<b>-1.68 (.70)</b>	<b>-2.34 (2.22)</b>
Emergency certification in subject	<b>.58 (.70)</b>	<b>.84 (1.33)</b>	<b>.27 (.31)</b>	<b>1.00 (1.53)</b>
Private school certification	<b>-1.26 (1.60)</b>	<b>-.64 (1.08)</b>	<b>-1.55 (1.88)</b>	<b>-.51 (.82)</b>
Not certified in subject	<b>-1.35 (1.99)</b>	<b>-.78 (1.07)</b>	<b>-1.67 (2.30)</b>	<b>-.63 (.84)</b>
Standard certification * admissions test	-	-	<b>-.20 (.71)</b>	<b>-.14 (.60)</b>
Probationary certification * admissions test	-	-	<b>2.82 (1.03)</b>	<b>2.40 (1.89)</b>
Standard certification * licensure test	-	-	<b>-.28 (1.12)</b>	<b>.39 (2.05)</b>
Probationary certification * licensure test	-	-	<b>1.06 (.45)</b>	<b>1.89 (1.51)</b>

# Summary of Our Results

- Students of teachers with no subject-specific training perform worse in math
- No statistically significant difference between student test scores of standard, probationary & emergency certified teachers
- No evidence that differences in state certification policies affect teacher performance as measured by student achievement

# Sample Questions from Praxis Exams

## Curriculum Test Questions

In which of the following theories is the influence of rewards most likely to be emphasized in explaining behavioral change?

- (A) Information-processing theory
- (B) Operant conditioning theory
- (C) Classical conditioning theory
- (D) Cognitive development theory

An 8 year old tries to ice-skate by moving her legs in the same way that she has done when roller-skating. Which of the following of Piaget's concepts of development does this behavior exemplify?

- (A) Accommodation
- (B) Assimilation
- (C) Reversibility
- (D) Egocentrism

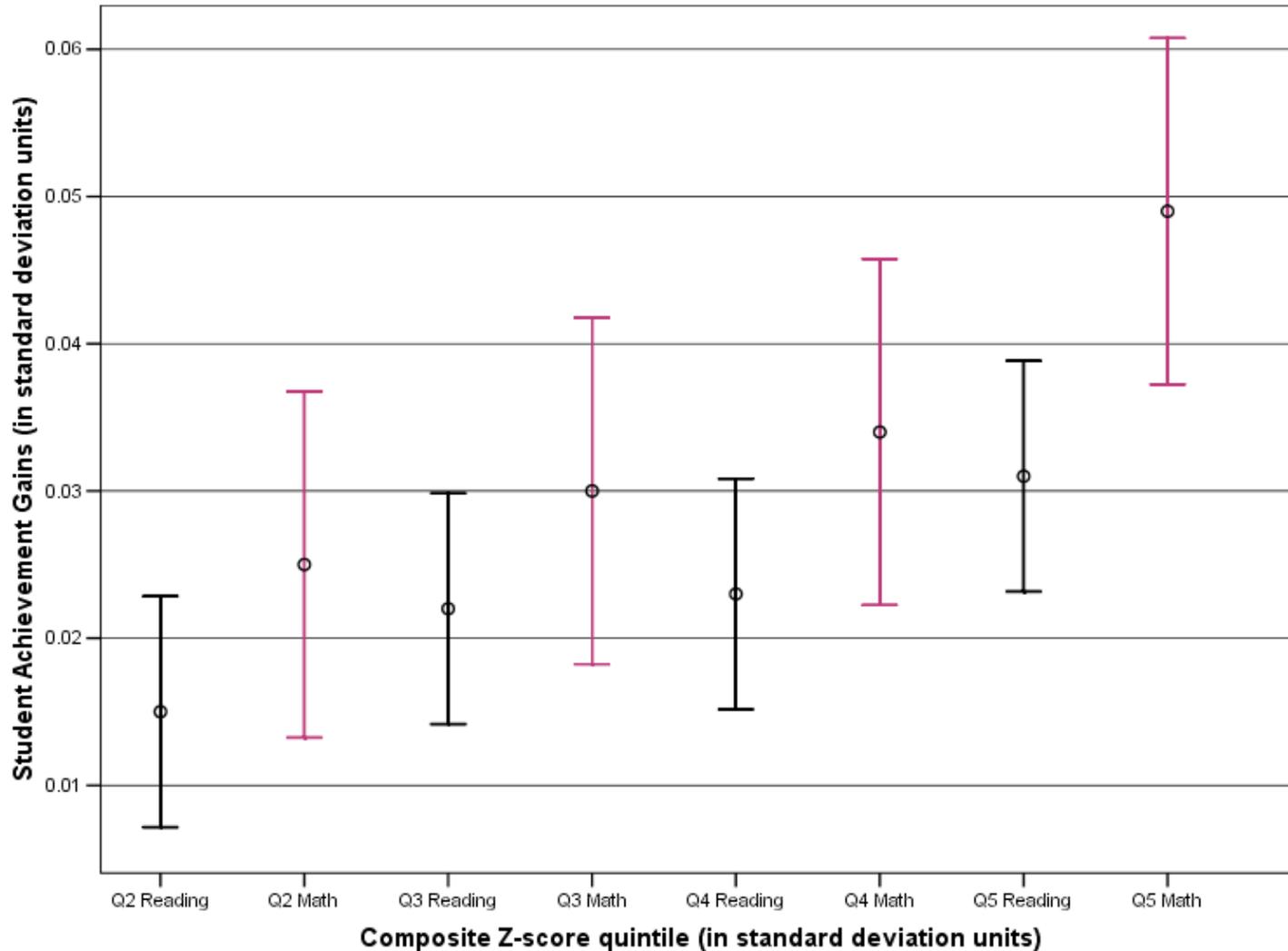
## Content Test Question

*Feb 3.  
Bank robbing  
The people were at the bank.  
A Robber stold a million  
dollers. That night was a  
terrible night.*

Describe two (2) ways the teacher could assist this student in the continued development of writing skills. The first example should include one or more activities that would build on the student's strengths in written expression. The second example should suggest a way for building on the student's strengths in the mechanics of writing.

# Graphical Depiction of Teacher Licensure-Test Performance

## Teacher Effectiveness Along Test Distribution



# Comparison of Teacher Effects in Math by Passing Status

QuickTime™ and a  
TIFF (Uncompressed) decompressor  
are needed to see this picture.