10 Reasons to Replace Midterms with Weekly Online Quizzes

Dave Pritchard, Cecil and Ida Green Prof. Physics

Byron Drury, Postdoc in RELATE

REsearch in Learning Assessing and Tutoring Effectively

1. Evidence Midweek Testing > Midweek Review

- **Learn Monday, Test Friday**

- **Wednesday has Review and Test, some words omitted**

- **Clearly Wednesday testing vs not testing determines Friday test score**
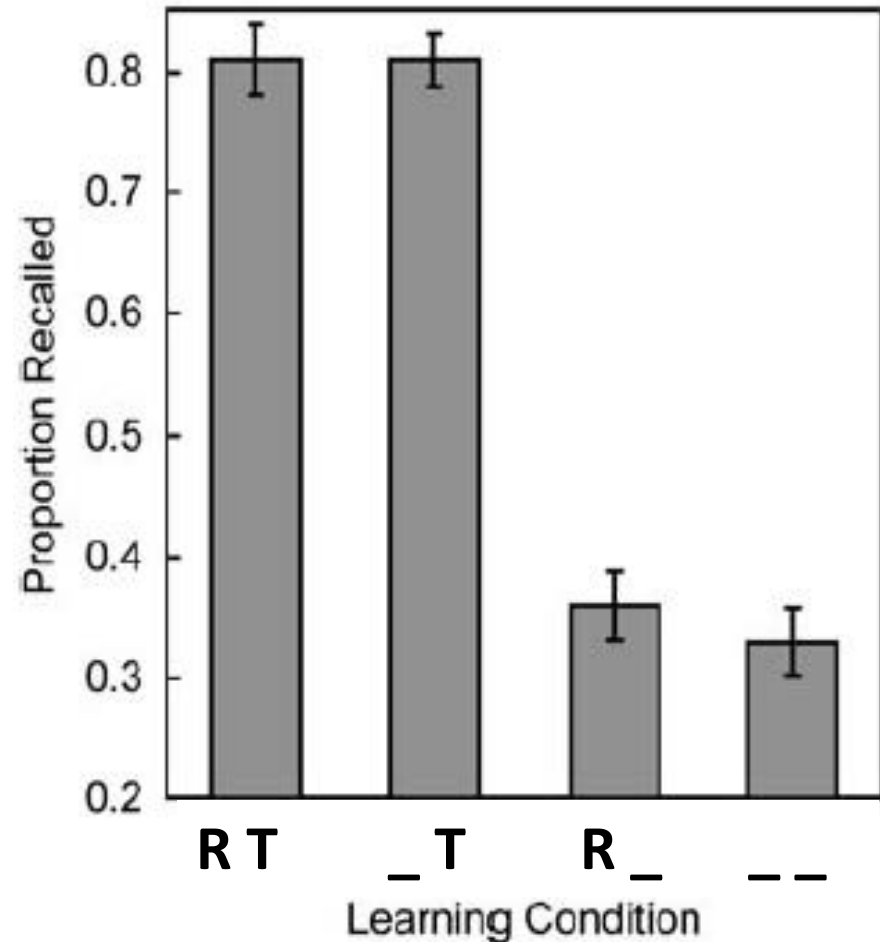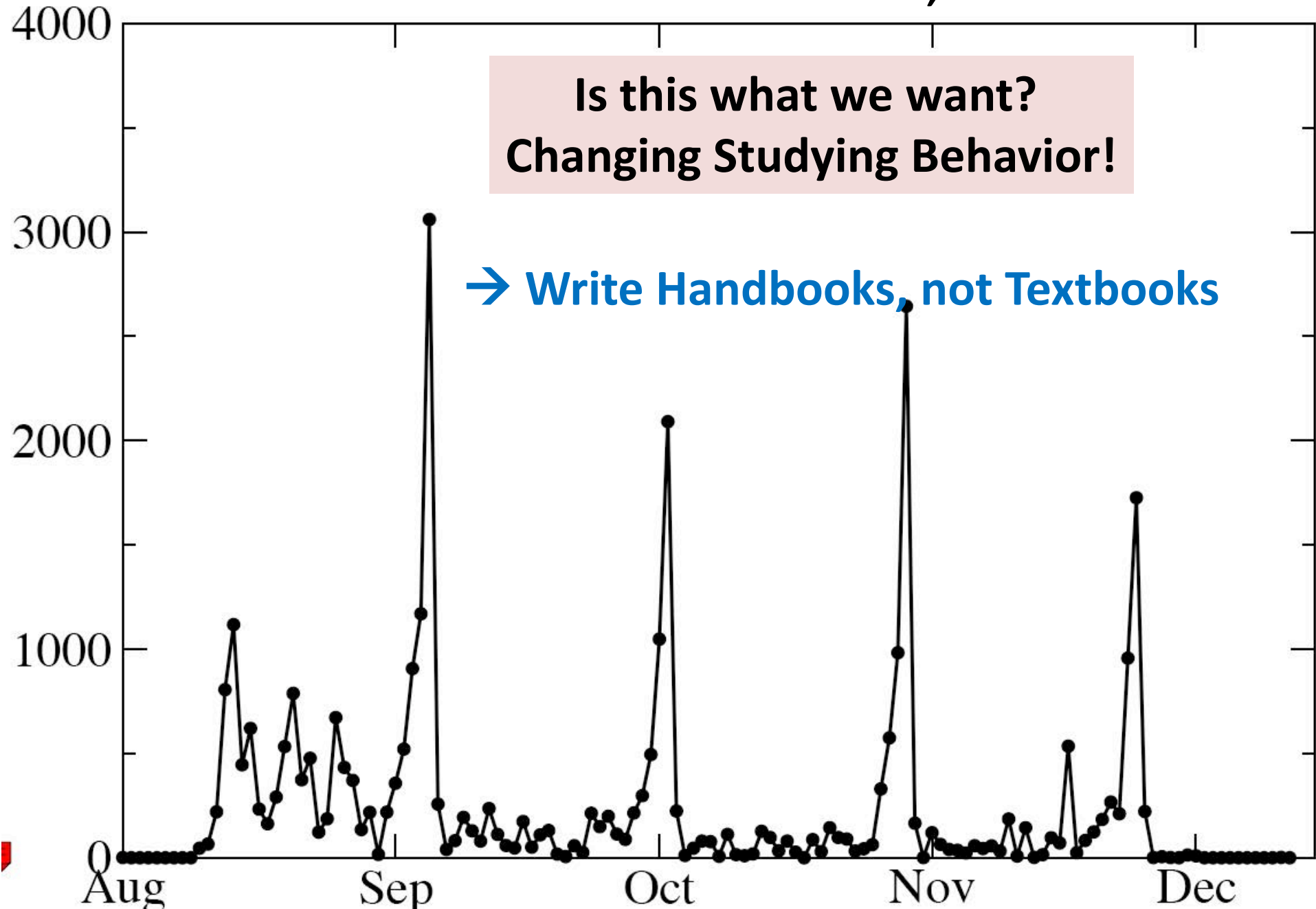
Karpicke & Roediger
Science 319, 966 (2008)

Fig. 2. Proportion recalled on the final test 1 week after learning. Error bars represent standard errors of the mean.

When do Students Read e-text; traditional?

Is this what we want?
Changing Studying Behavior!

→ **Write Handbooks, not Textbooks**

When Do Students Study?

Question? "How often do you Study, by which I mean review & reflect & do extra problems vs. just doing the weekly p-sets?
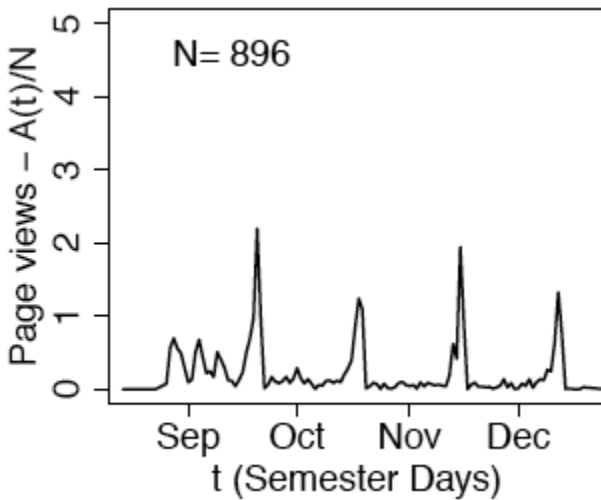
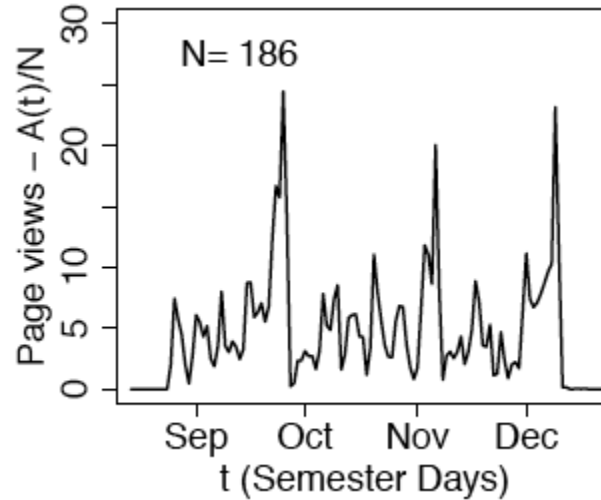"I just study for exams"

"Only if I have a test"

"Are you kidding?"

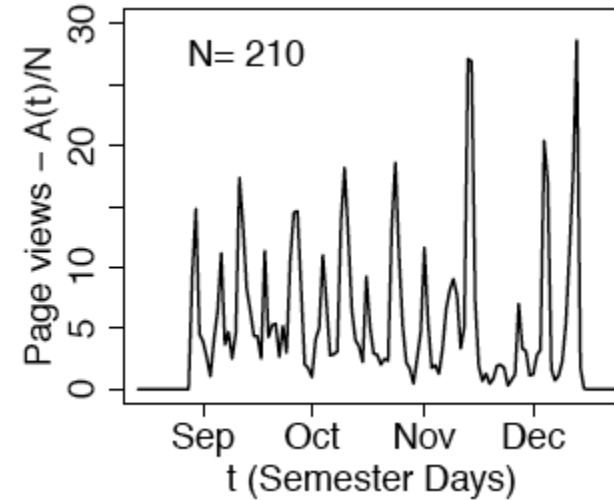# 2 text reading depends on Course Structure

Traditional   **Flipped**   **Reformed & Biweekly**

15>15   55>55   77>77

Traditional: Lectures, etext, weekly homework

Reformed: Flipped Classroom with preclass HW

Reformed & Biweekly: Same plus quizzes

Copying Homework Influences Final Exam Scores

Palazzo, D. et. al. Phys. Rev. ST Phys. Educ. Res. vol. 6, (2010), p. 010104

Legend:
- Copy rate < 0.1
- 0.1 ≤ Copy rate < 0.3
- 0.3 ≤ Copy rate < 0.5
- Copy rate ≥ 0.5

Exam t-scores

Letter grade divisions: A | B, B | C, C | D

Net of 1.1 std dev for about 60% copying implies ~ 2.0 std. dev. effect size for *no copying* vs. *all copying*

Pre test   Exam 1   Exam 2   Exam 3   Final exam
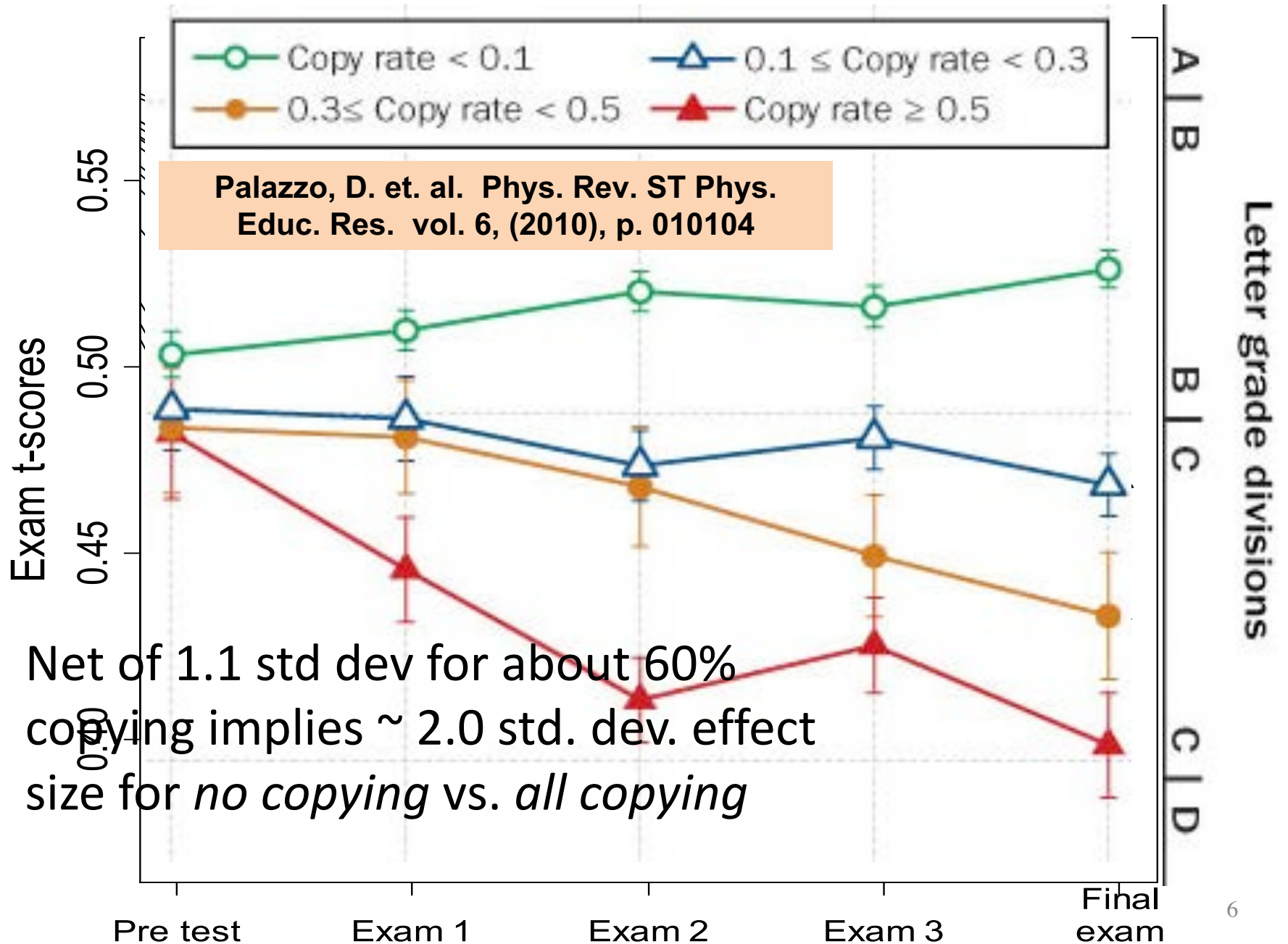
Homework Copying is Best Predictor of Final

final exam score =

$-0.47C$	*Copying online HW*
$+0.26EX1$	*First Midterm EXam*
$+0.26S$	*Score on HW*
$+0.20D$	*Diagnostic Test*

→ **One might expect that *decreasing* the use of unsupervised answer-sharing forums would improve final exam scores**

3. Monitored vs. Unmonitored Discussion Forum

- **Students switched to using discussion forum monitored by TA's vs. ones that allow posting of answers**

Laverty et. al. *Want to Reduce Guessing and Cheating While Making Students Happier?* ***Give More Exams!***
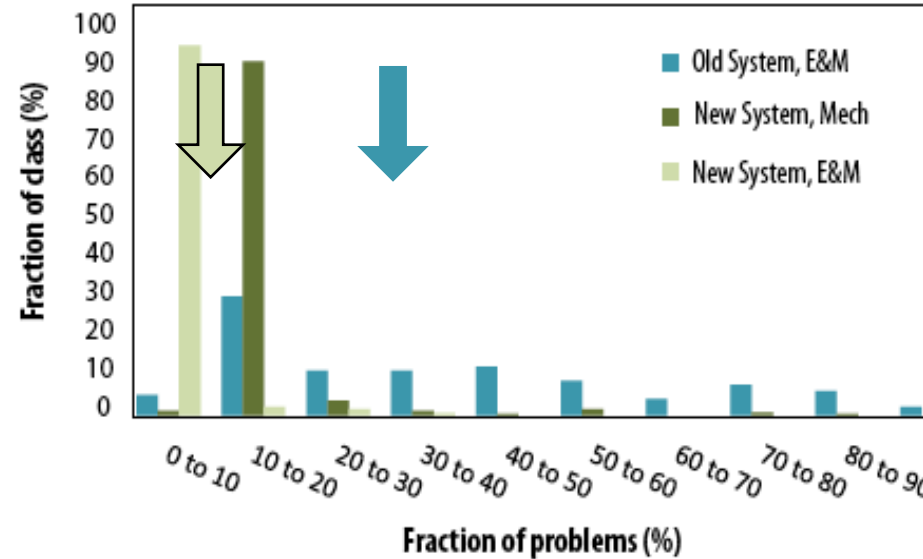

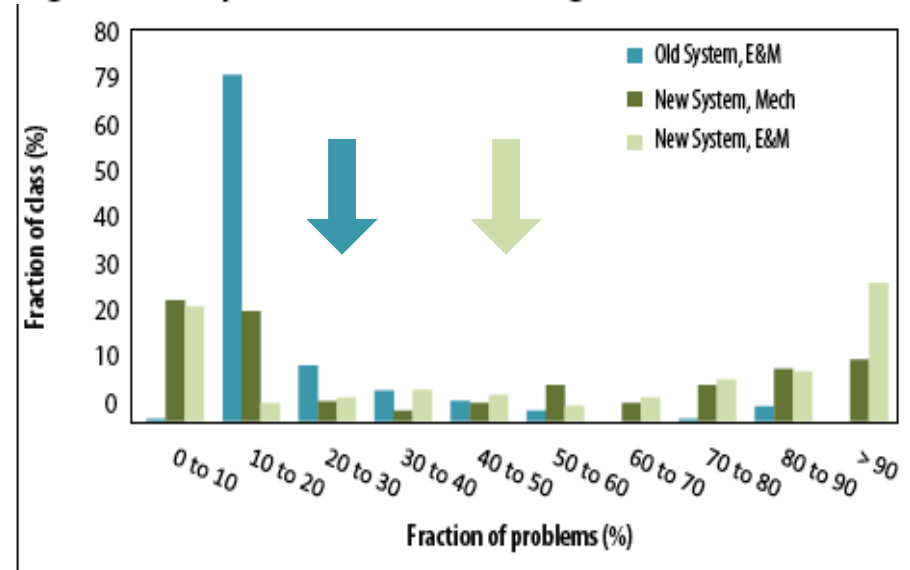Fig. 2. Self-reported use of "cheating sites."

Fig. 3. Self-reported use of sanctioned internal discussion board.

Weekly Quizzing Should Increase Learning
Summary

- Rodighier: Quizzing Improves Learning

- Students Study Every Week (not just do HW)

- Students don't use homework copying forums
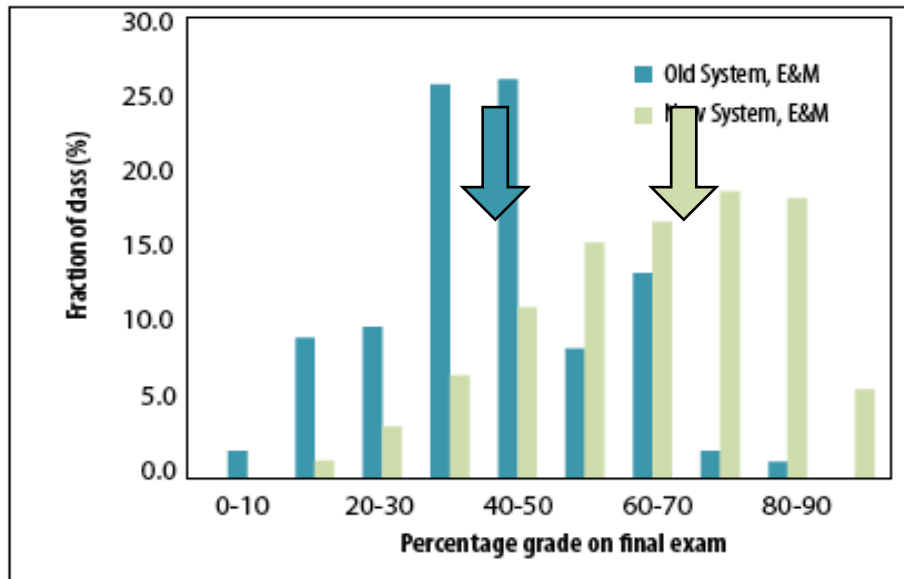
Indeed, it Dramatically Increases Final

Fig. 6. Final exam grade distribution.

- **Final average shifted from ~43% to ~68%.**

- **Large reduction in D & F grades**

4. Students liked frequent quizzes

[Students] "seemed unhappy with the [testing] system when they first heard about it", but gave very much more positive opinions after trying it.
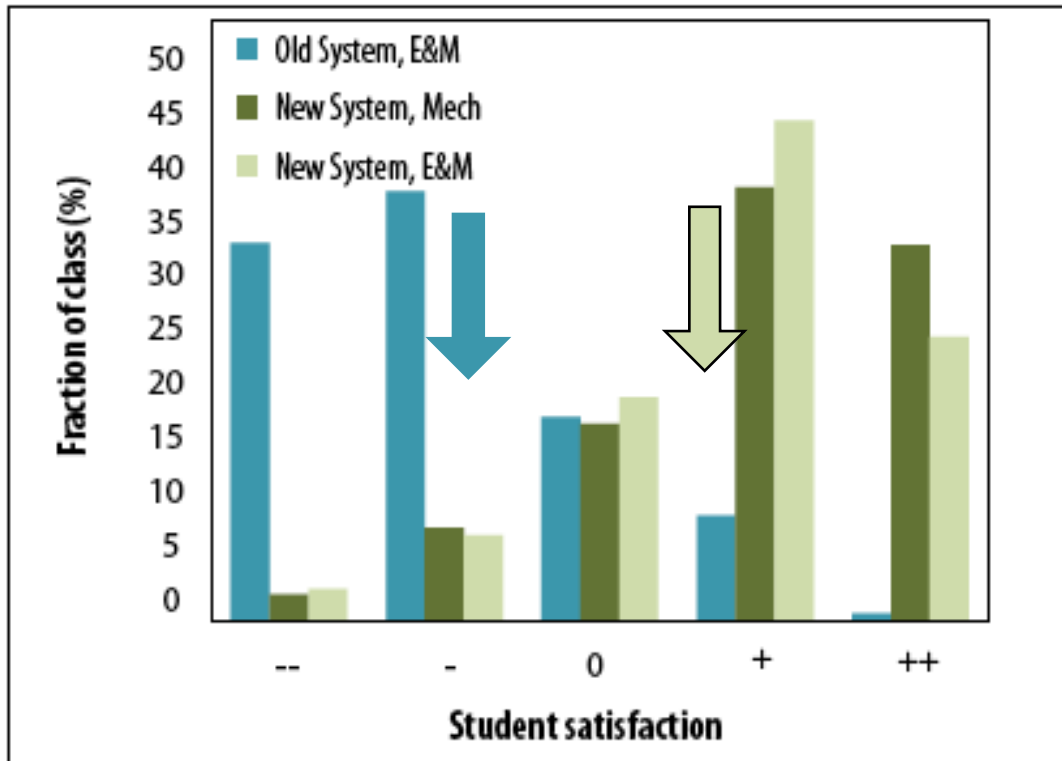

Fig. 5. Student satisfaction with exam schedule.

Online vs. OnPaper Quiz Experiment Spring 2017 8.011

Gave weekly quizzes in Friday's class

½ online and ½ on-paper

Students could allocate time as they desired;

   online/on-paper correlation: r~0.3 (week-by-week)

                                 r~0.7 (average over weeks)

Online Quiz: Concept vs. Calculation

A blend of question types and response formats:

> "conceptual:" multiple choice, checkbox

> "calculational:" symbolic input, numerical

Average ~3 min/question, but with large variation

> Conceptual shorter, Calculational longer

> ~10 questions for ½ hour quiz

Administered in-class

Students bring computers, ODL supplies loaners (~6%)

Quiz Grading

Many online questions allow multiple attempts;

We can measure student skill better if we give partial credit for correct answers submitted after initial incorrect attempts

Different problem types give different information about student knowledge
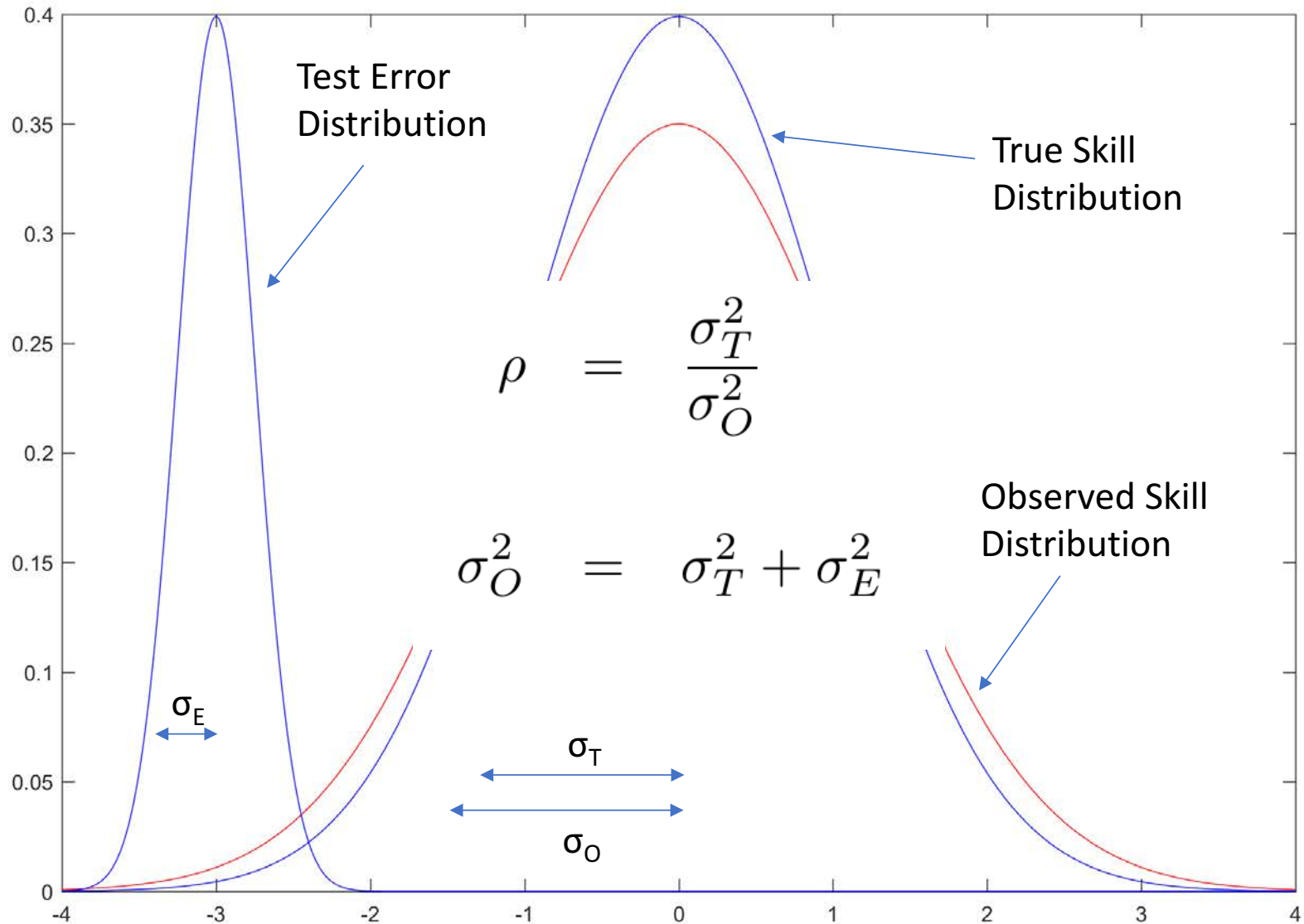
We explored a family of grading schemes:

$$s_{adj}(\theta, P_M, P_O) = \cos^2(\theta)\left[(1 - P_M)s_{M,1^{st}} + P_M s_{M,ev.}\right]$$
$$+ \sin^2(\theta)\left[(1 - P_O)s_{O,1^{st}} + P_O s_{O,ev.}\right]$$

θ controls relative weight of MC and OR questions

$P_{M/O}$ are amount of credit awarded for correct answers given *after* an incorrect attempt on MC/OR

Test Reliability

Test Error Distribution

True Skill Distribution

$$\rho = \frac{\sigma_T^2}{\sigma_O^2}$$

Observed Skill Distribution

$$\sigma_O^2 = \sigma_T^2 + \sigma_E^2$$

σ_E

σ_T

σ_O

Quiz Week to Week Self-Consistency

Need to quantify effects of testing error

use *z-scores* to normalize for week-to-week variation in difficulty

$$z = \frac{score - quiz\ mean}{\sigma_{quiz}}$$

Quantify the self consistency with Cronbach's α

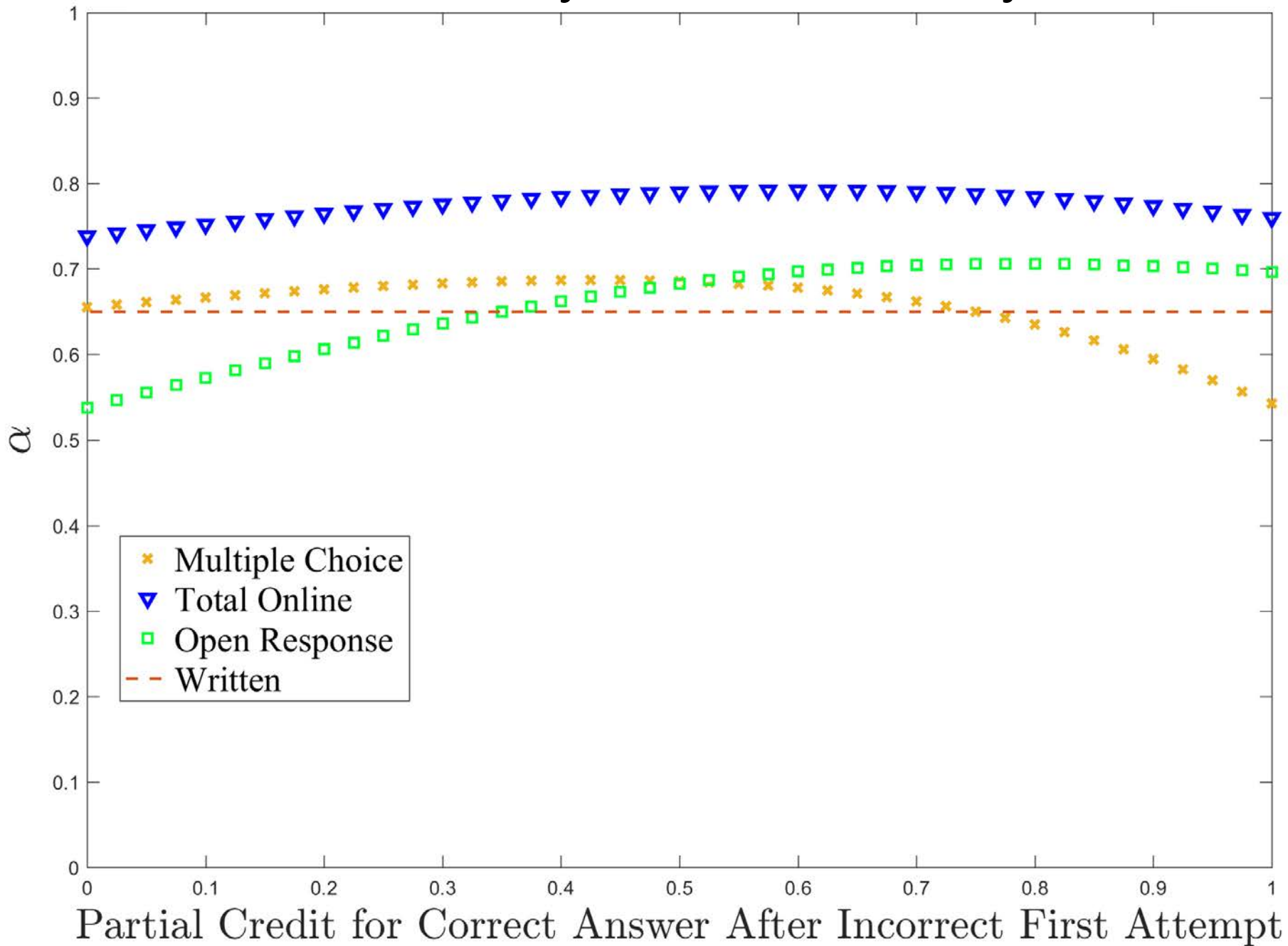$$\alpha = \frac{K\bar{c}}{(\bar{v} + (K - 1)\bar{c})}$$

K is # of tests, $\bar{v}$ is mean test variance, $\bar{c}$ is mean test covariance

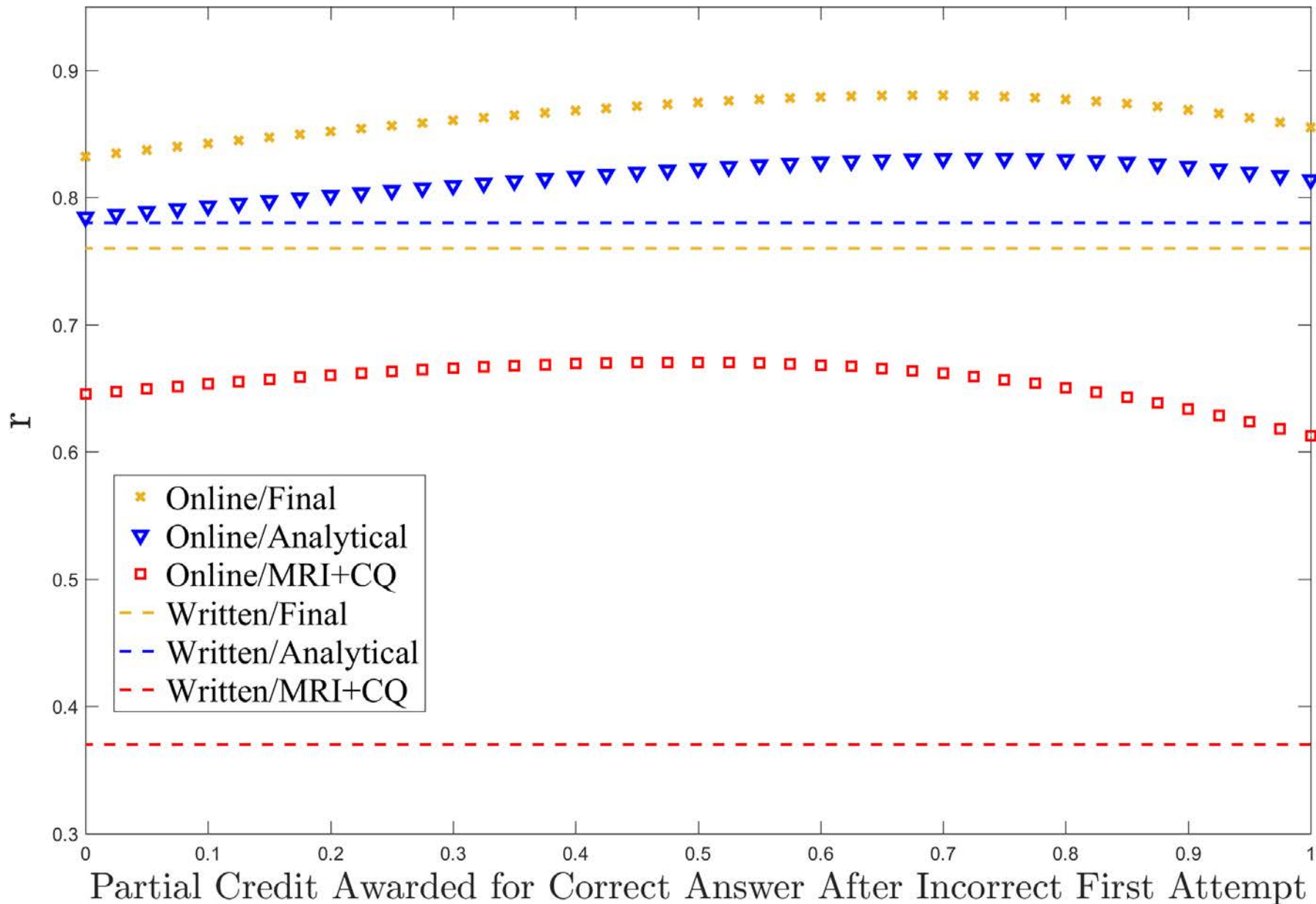Quantifies the extent to which the quizzes measure the same underlying construct

Ranges from 0 (no correlation) to 1 (perfect correlation)

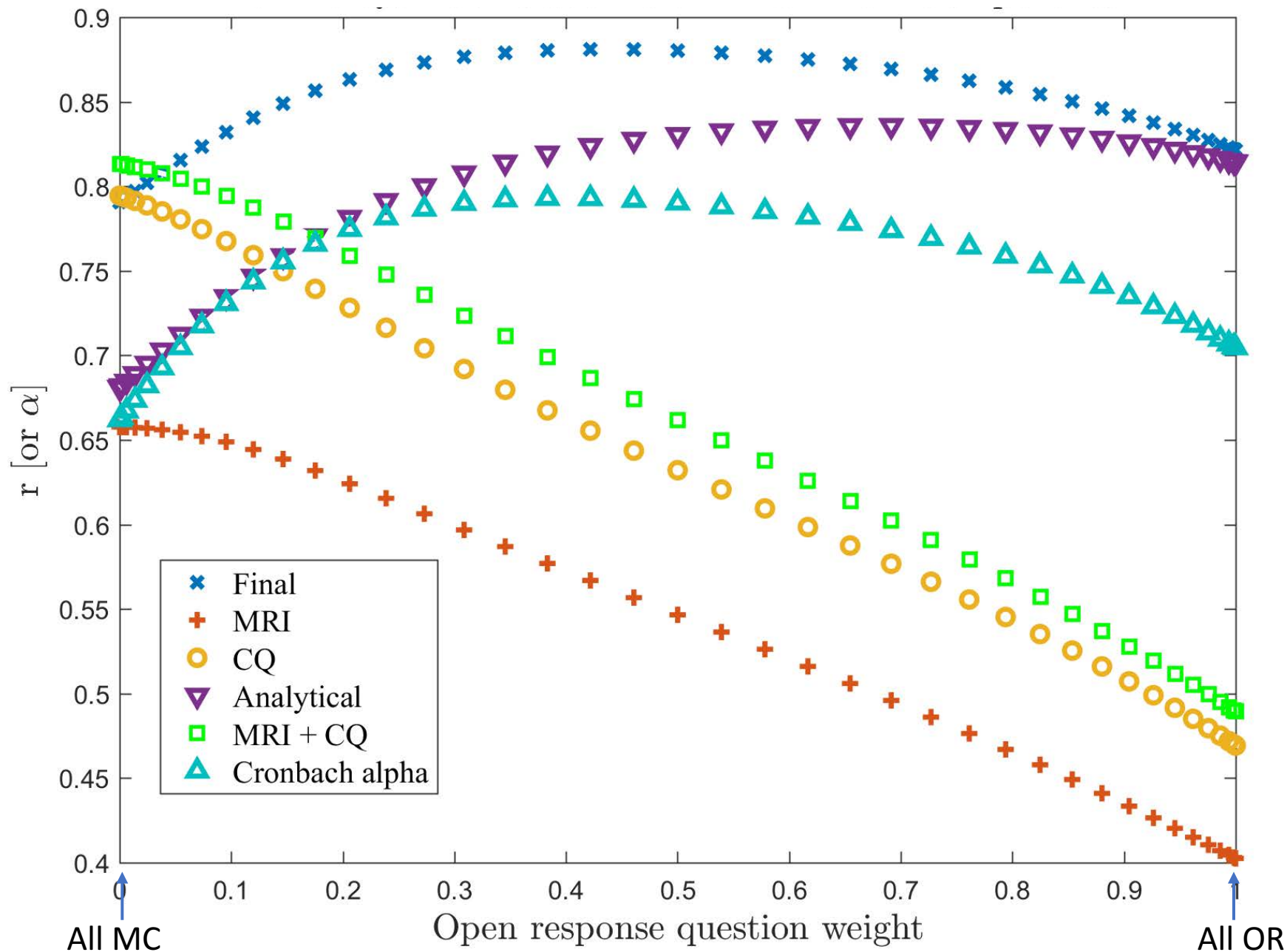Rule of thumb: 0.7+ is 'acceptable,' 0.8+ is 'good'

Quiz Weekly Self-Consistency

Quiz Correlation with Final

Different Question Formats test Different Knowledge?

Summary

Scheme		Cronbach's α	Corr. w/ Final	p-value
Written		0.65	0.76	0.00110
Online	First attempt	0.74	0.83	0.00012
	70% credit	0.78	0.86	0.00004
	Any attempt	0.76	0.86	0.00005
	Optimized	0.80	0.88	0.00002

Online quizzes are more consistent than written and predict performance on the final better

Weekly Online Quizzing Summarized

Fairer, and predicts every part of Final better than rubric-graded symbolic response quizzes, eliminates grading error

Facilitates weekly quizzes: research→ less dishonesty, far better learning

Less faculty time spent on grading/testing

Facilitates education research

Year-to-year and student-to-student grading consistency → evaluate teaching by results

Discussion

- Difficulties with student acceptance?

- Quiz Security vs. Releasing Quizzes
    - Is knowing in real time enough?
    - Better than paper quizzes in fraternity Bibles?
    - Online offers easy checks

- Implementing in multi-section TEAL subjects?