Pre-instruction diagnostic tests can help predict probability of obtaining high or low course grades in introductory physics

David E. Meltzer[1] and Dakota H. King[2]

[1]Arizona State University
[2]University of Arizona College of Veterinary Medicine

Supported in part by NSF DUE #1504986 and #1914712

30 years ago today (+1 day): My first AAPT talk

Session FB: Contributed Session on Topics in Introductory Physics—College Level, Part I (Also see Part II–GC, January 19, 8:30 a.m.)

Wednesday, January 18, 1995; 1:15 p.m.–5:15 p.m.

Salons 11 & 12

Presiding: LARRY BADAR, Dept. of Physics, Case Western Reserve Univ., 10900 Euclid Ave., Cleveland, OH 44106-7079; 216-368-8779

[illegible]

FB12 4:00 A Pilot Project for an Elementary Phys-[illegible]

David E. Meltzer, Dept. of Chemistry and Physics,
Southeastern Louisiana Univ., Hammond, LA 70402;
504-549-2158, FAX: 504-549-5126, dmeltzer@selu.edu

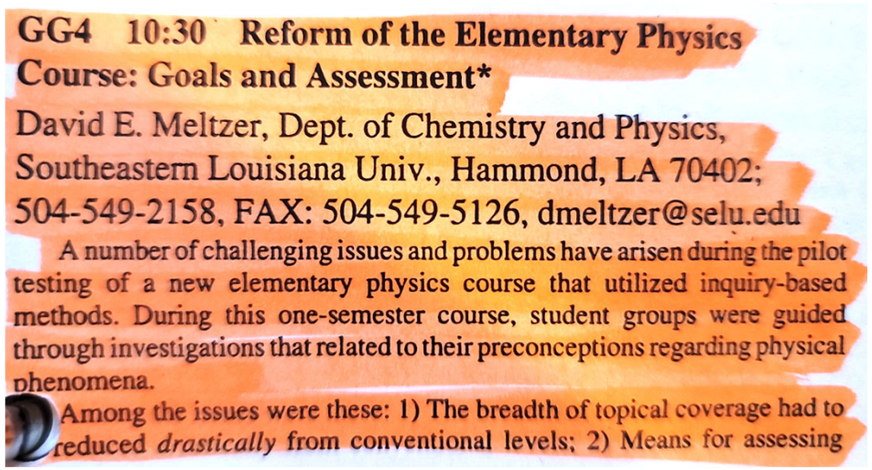
"A Pilot Project for an Elementary Physics Course Based on Guided Inquiry, with the Theme of 'Energy'"

One year later, I reported some concerning findings…

Session GG: Invited and Contributed Session on Reforms in High School Physics Education (Sponsored by the Committee on Physics in High Schools)

Thursday, January 18, 1996; 9:00 a.m.–11:00 a.m.

Bonanza B



GG4 10:30 Reform of the Elementary Physics Course: Goals and Assessment*

David E. Meltzer, Dept. of Chemistry and Physics, Southeastern Louisiana Univ., Hammond, LA 70402; 504-549-2158, FAX: 504-549-5126, dmeltzer@selu.edu

A number of challenging issues and problems have arisen during the pilot testing of a new elementary physics course that utilized inquiry-based methods. During this one-semester course, student groups were guided through investigations that related to their preconceptions regarding physical phenomena.

Among the issues were these: 1) The breadth of topical coverage had to reduced *drastically* from conventional levels; 2) Means for assessing

"A number of challenging issues and problems have arisen during the pilot testing of a new elementary physics course…"

→ Typical students have very weak academic preparation

- poor algebra skills
- difficulty with proportional reasoning
- very weak graphing skills
- little experience or ability in conceptual reasoning

- A great deal of time had to be spent on "remedial" work, such as graphing skills and rate/time/distance problems

- There was sustained difficulty with rudimentary algebra and proportional reasoning

- Students had great difficulty with abstract concepts such as instantaneous velocity, acceleration, and energy

Are effects of these same issues evident in the general physics course?

- We administered a variety of diagnostic pretests to students in introductory general physics courses
- We assessed the degree to which performance on the pretests is associated with students' final grades in physics

Assessment Pretests

- Diagnostic pretest covering pre-college mathematics (“Math”)
  - calculators allowed
- Pre-instruction tests of scientific reasoning skill and physics concept knowledge:
  - Lawson Test of Scientific Reasoning (“Lawson”)
  - Force Concept Inventory (FCI)
- Why do this? Perhaps ultimately we can offer special assistance to those students who need it most.

Sample Description

- 31 introductory physics classes from 4 universities, 8 different instructors; over 2000 total students.
- Instruction in most classes was “non-traditional,” generally highly interactive using research-based instructional materials and methods

Acknowledgments

- Diagnostic data have been provided by (among others):

  Vince Coletta (Loyola Marymount University)

  Steven Pollock (University of Colorado, Boulder)

  Christopher Varney (University of West Florida)

Course and Institution Code

Alg-1: Algebra-based course, first semester

Alg-2: Algebra-based course, second semester

Calc-1: Calculus-based course, first semester

Calc-2: Calculus-based course, second semester

ASU-P: Arizona State University, Polytechnic campus

ASU-T: Arizona State University, Tempe campus

LMU: Loyola Marymount University

UWF: University of West Florida

CU: University of Colorado, Boulder

Comparing probabilities of high and low grades

- What is the probability of a student with a high score on a pre-instruction assessment getting a high grade in the class?
- How does that compare to a low-scoring student's probability of getting a high grade?

(and, same questions for probabilities of getting a *low* grade)

Consistent result:

High (top-quartile) scorers on *any one* of the diagnostic pretests were much more likely to get high (top-quartile) grades and much less likely to get low (bottom-quartile) grades than were low scorers.

But how *much* more (or less) likely?…

- High scorers on any one of the pretests were much more likely (400-600%) to receive high grades than were low scorers.
- High scorers were much *less* likely (20-30%) to receive low grades than were low scorers.
- This general pattern held for 113 out of 116 comparisons (97%) and for all four universities, although the quantitative range was large.

High-Grade Probability vs. Lawson Pretest Score

Course	Campus	*N*	Top-quartile Lawson: % with top-quartile grades	Bottom-quartile Lawson: % with top-quartile grades	High-grade odds ratio
Alg-1 2021a	ASU-P	37	49%	11%	4.3
Alg-1 2021b	ASU-P	36	41%	11%	3.7
Alg-1 2022a	ASU-P	41	49%	10%	5.0
Alg-1 2022b	ASU-P	53	58%	10%	5.8
Alg-1 2023a	ASU-P	36	39%	33%	1.2
Alg-1 2023b	ASU-P	43	55%	10%	5.5
Alg-2 2022	ASU-P	66	52%	4%	11.9
Alg-2 2023	ASU-P	76	51%	16%	3.2
Alg-2 2024	ASU-P	90	41%	5%	8.0
Alg-1 2005	CU	469	45%	8%	5.5
Calc-2 2007	CU	276	57%	8%	6.9
Alg-1 2007	LMU	24	50%	0%	[undefined]
Alg-1 2009	LMU	51	34%	11%	3.2
Alg-1 2011	LMU	57	53%	18%	2.9
Alg-1 2012	LMU	44	64%	6%	10.5
Alg-1 2013	LMU	30	53%	12%	4.6
Alg-1 2014	LMU	33	61%	0%	[undefined]
Alg-1 2015	LMU	24	63%	0%	[undefined]
Alg-1 2016	LMU	35	41%	0%	[undefined]
Alg-1 2018	LMU	47	54%	9%	6.3
Alg-1 2021	LMU	27	44%	0%	[undefined]
AVERAGE (unweighted)		***[1595]***	**50%**	**9%**	**5.8**

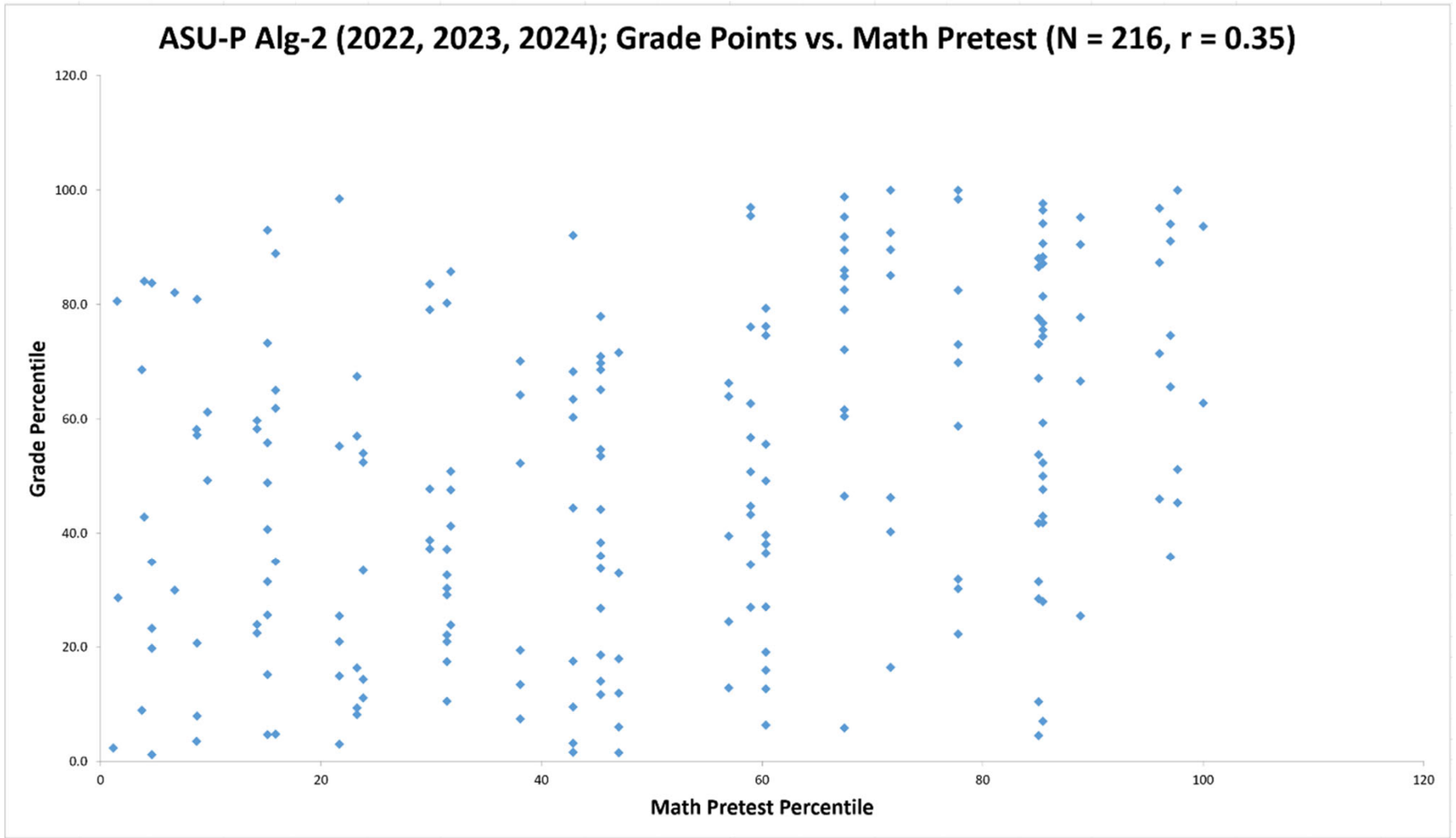
High scorers almost 6 times as likely to get top-quartile grades as low scorers

Low-Grade Probability vs. Lawson Pretest Score

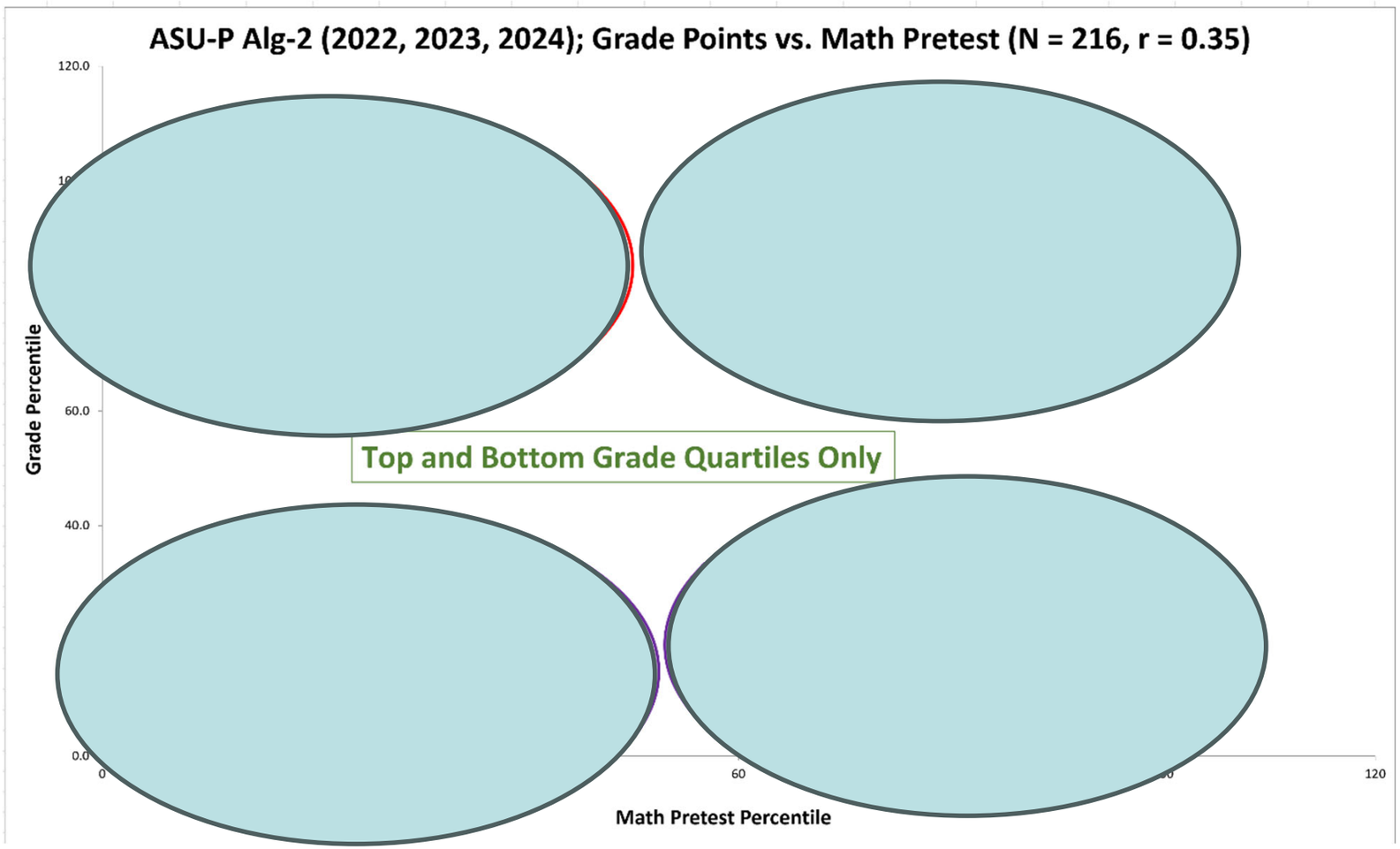
Course	Campus	*N*	Top-quartile Lawson: % with bottom-quartile grades	Bottom-quartile Lawson: % with bottom-quartile grades	Low-grade odds ratio
Alg-1 2021a	ASU-P	37	6%	44%	7.2
Alg-1 2021b	ASU-P	36	11%	47%	4.2
Alg-1 2022a	ASU-P	41	15%	28%	1.9
Alg-1 2022b	ASU-P	53	15%	45%	3.0
Alg-1 2023a	ASU-P	36	14%	36%	2.6
Alg-1 2023b	ASU-P	43	8%	50%	6.7
Alg-2 2022	ASU-P	66	12%	25%	2.1
Alg-2 2023	ASU-P	76	11%	28%	2.7
Alg-2 2024	ASU-P	90	10%	36%	3.6
Alg-1 2005	CU	469	10%	42%	4.4
Calc-2 2007	CU	276	12%	44%	3.8
Alg-1 2007	LMU	24	0%	58%	[undefined]
Alg-1 2009	LMU	51	5%	48%	10.4
Alg-1 2011	LMU	57	15%	46%	3.0
Alg-1 2012	LMU	44	9%	27%	3.0
Alg-1 2013	***LMU***	***30***	***27%***	***12%***	***0.4***
Alg-1 2014	LMU	33	0%	68%	[undefined]
Alg-1 2015	LMU	24	0%	75%	[undefined]
Alg-1 2016	LMU	35	11%	46%	4.0
Alg-1 2018	LMU	47	16%	42%	2.7
Alg-1 2021	LMU	27	0%	89%	[undefined]
AVERAGE (unweighted)		***[1595]***	**10%**	**45%**	**4.5**

Low scorers 4.5 times as likely to get bottom-quartile grades as high scorers

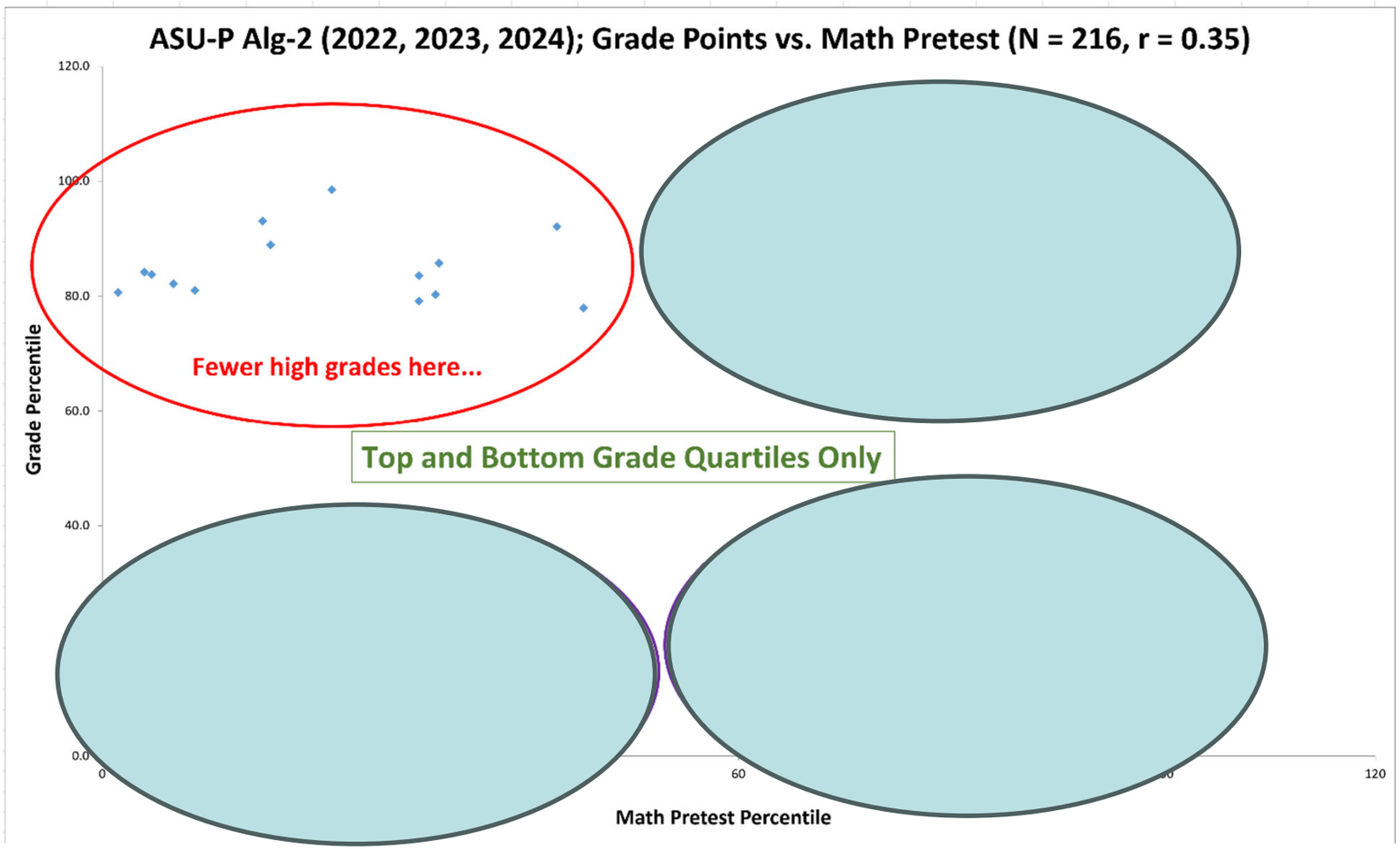
Grades vs. Math Pretest Score for “Combined” Sample



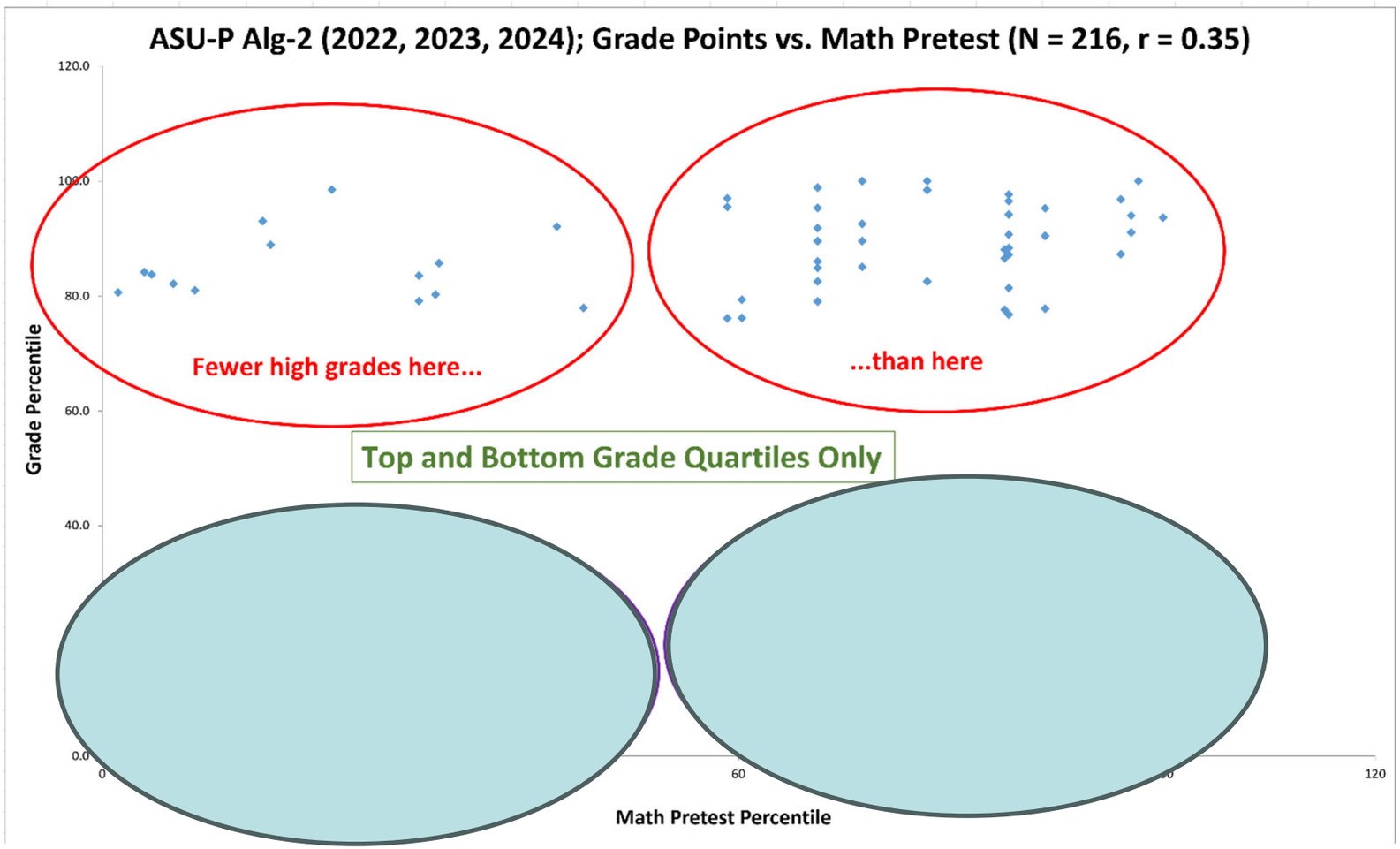
Grades vs. Math Pretest Score for “Combined” Sample



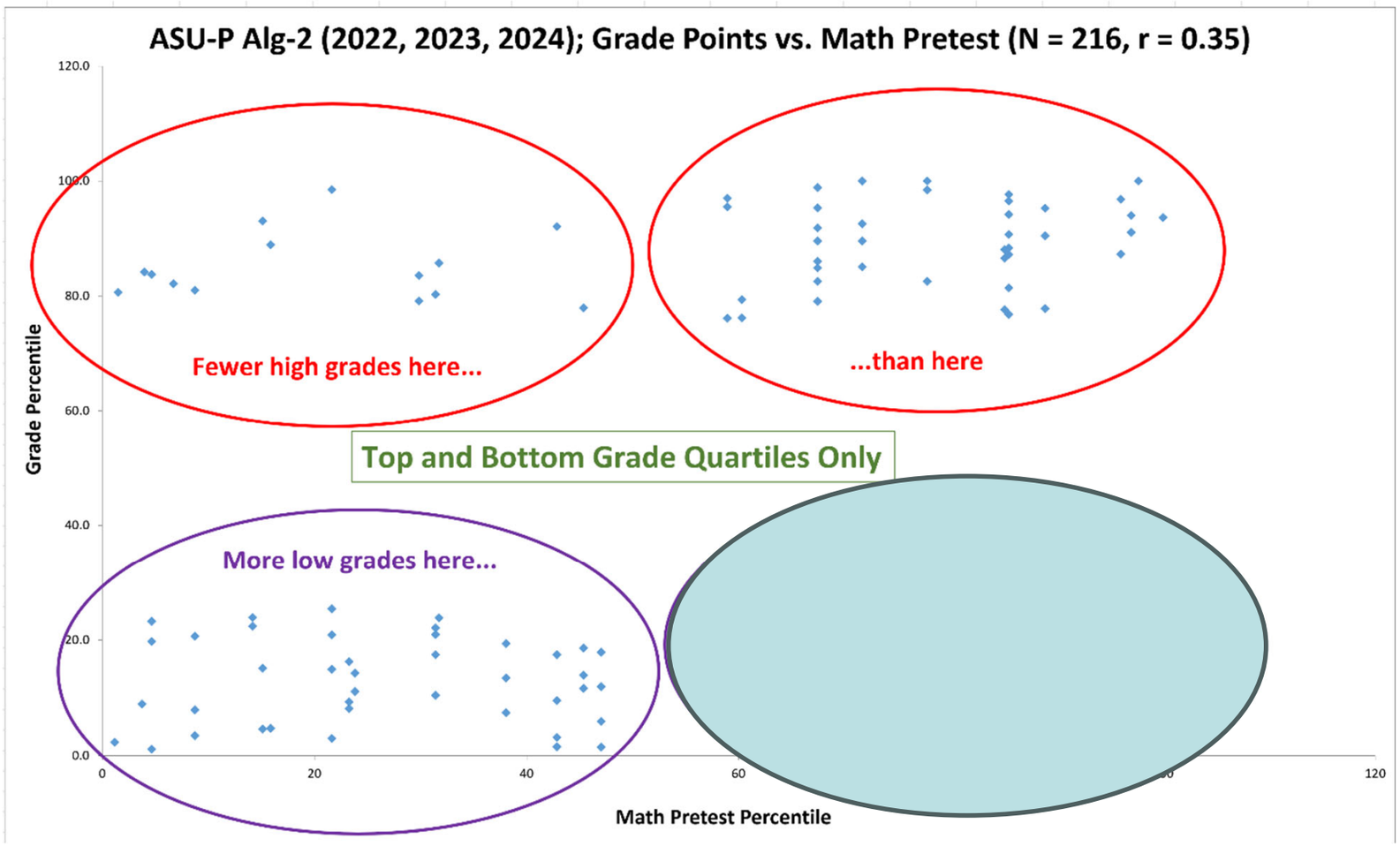
Grades vs. Math Pretest Score for “Combined” Sample



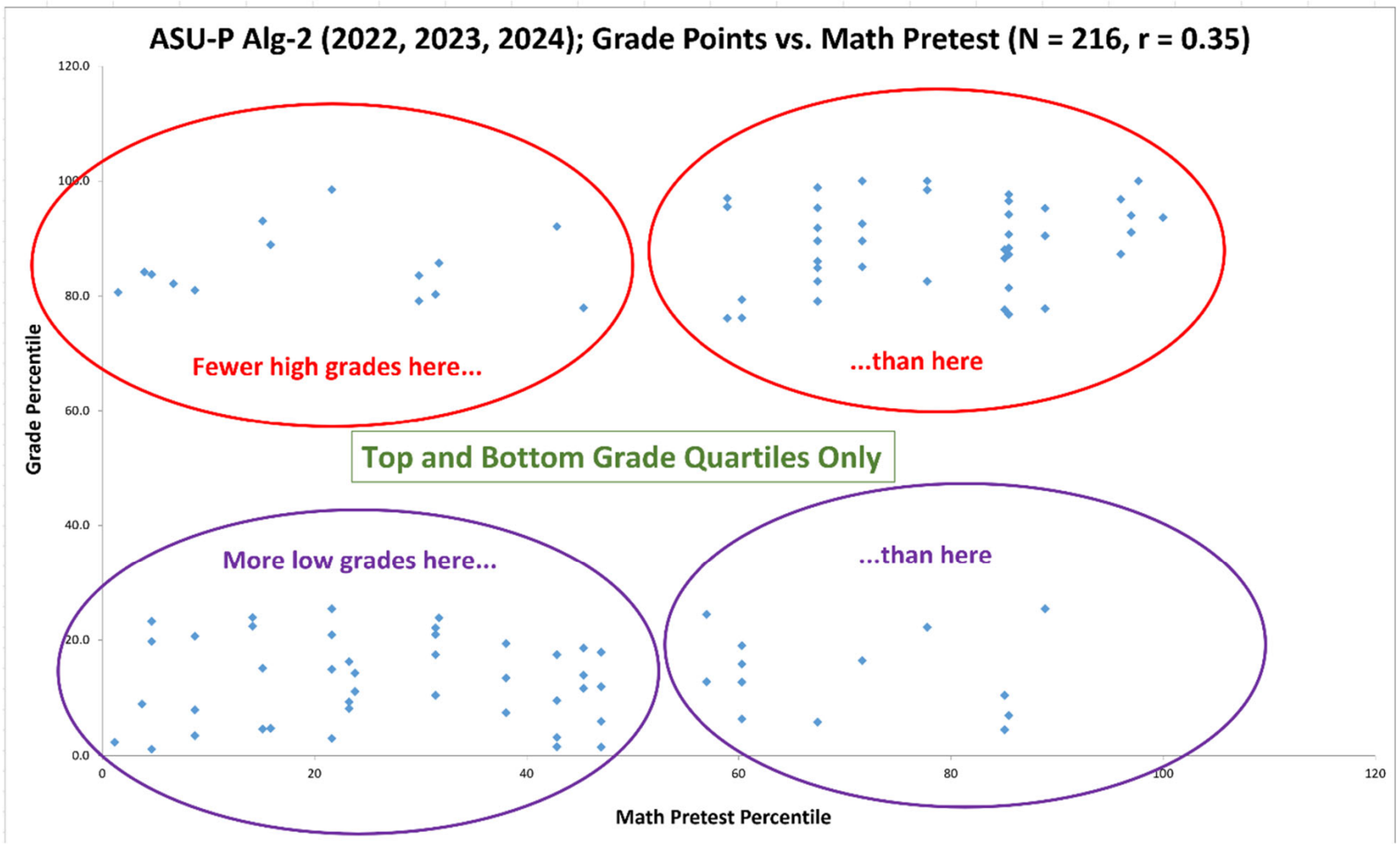
Grades vs. Math Pretest Score for “Combined” Sample



Grades vs. Math Pretest Score for “Combined” Sample



Grades vs. Math Pretest Score for “Combined” Sample

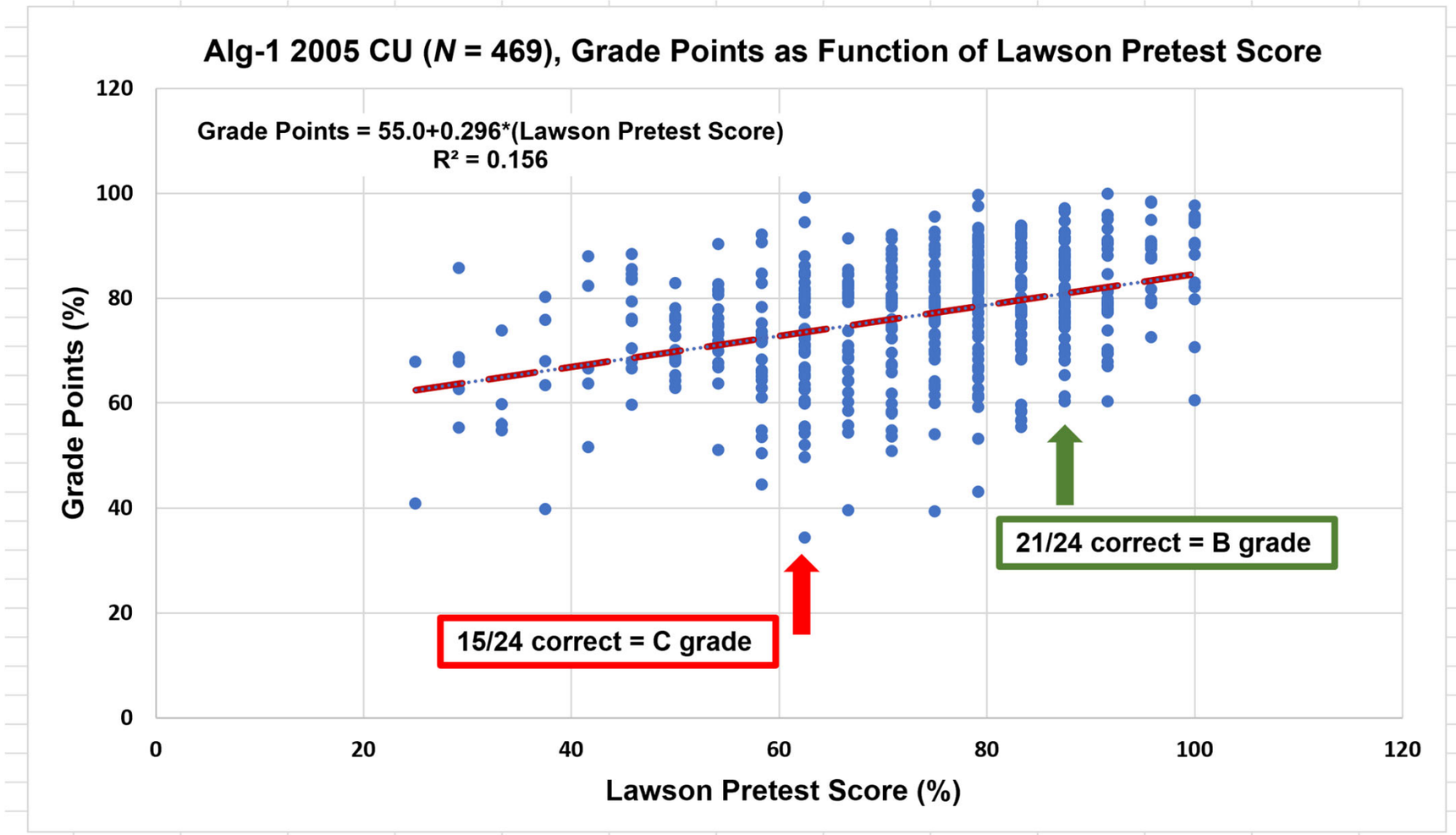


Grades vs. Lawson Pretest Score for Alg-1 2005 CU Sample

For Alg-1 CU, our largest sample (N = 469):

How do students' median course grades vary depending on their pretest scores on the Lawson test of scientific reasoning?

Grades vs. Lawson Pretest Score for Alg-1 2005 CU Sample



Some “answers” to relevant questions

- Which, if any, of the diagnostic pretests is most predictive of students’ performance? *Varies with the course*
- Does using multiple predictor variables offer greater predictive power than using just one of them? *Yes, sometimes*
- Does better performance on one predictor variable indicate that another variable is more (or less) predictive? (This would be an “interaction” effect.) *Maybe*

  [if student gets a high score on one pretest, score on the other pretest adds predictive power; if low score on one pretest, other pretest is not very predictive]

Prediction by Salehi et al. (2019)* based on their theoretical model and empirical observations in similar courses:

"These …[model-fit] values may seem modest to some, but they have career-altering implications for students who are poorly prepared….for [a typical value of the model-fitting parameter]…a student who comes in with preparation in the bottom quartile has about a factor of 4 higher probability of being in the bottom quartile of the grade distribution than a student who starts the course in the upper quartile of preparation. If one considers bottom quartile exam scores as failing, this means that poorly prepared students are 4 times more likely to fail their physics 1 final exams than peers with good incoming preparation." [Salehi et al. (2019), p. 020114-6]

Our results: "poorly prepared" students (i.e., low scorers) are 2-5 times more likely to get bottom-quartile course grades than peers with "good" preparation (i.e., high scorers)

*Shima Salehi, Eric Burkholder, G. Peter Lepage, Steven Pollock, and Carl Wieman, "Demographic gaps or preparation gaps?: The large impact of incoming preparation on performance of students in introductory physics," Phys. Rev. Phys. Educ. Res. **15**(2), 020114 (2019).

Summary

- Regardless of group *probabilities*, the course performance outcome for any *individual* student remains highly uncertain and depends on many factors.
- **It is reasonable to acknowledge that the course performance expected for the *group* of low-scorers on these pretests must be very different from that expected for the high-scorers.**
- Question: Can these findings be used to offer modified or supplemental instruction for the more “at-risk” group to improve their course outcomes?