Development And Analysis Of A Basic Proof Skills Test

Sandra Merchant & Andrew Rechnitzer
Carl Wieman Science Education Initiative
University Of British Columbia

Motivation

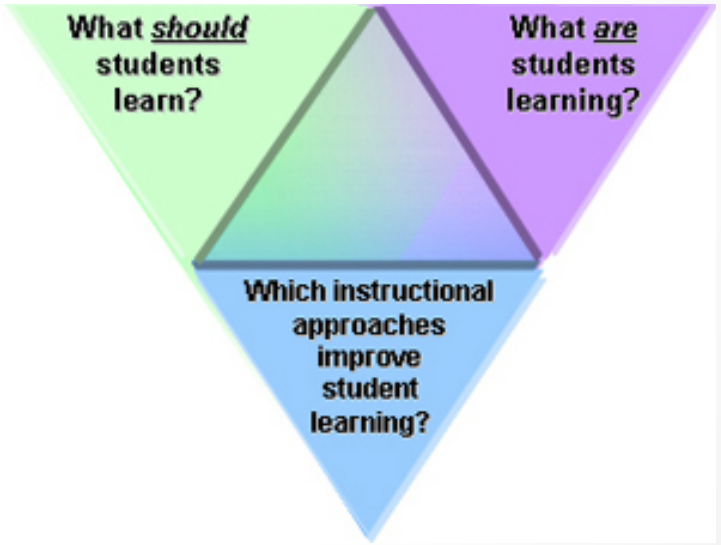
- MATH 220: "Mathematical Proof"
- Typical "transition-to-proof" course:
  - "Sets and functions; induction; cardinality; properties of the real numbers; sequences, series, and limits. Logic, structure, style and clarity of proofs emphasized throughout"
- Gateway to upper-level math
- Many factors led to CWSEI involvement
  - High failure rates
  - Instructor dissatisfaction with learning outcomes and teaching experiences
  - Reputation with students

The CWSEI Mandate:

Achieving the most effective, evidence-based science education

The Process:

1. Determine what students should learn
2. Measure what students are actually learning
3. Implement research-based instructional approaches to improve learning (and measure the result)
4. Disseminate and adopt what works



CWSEI Goals:

- Improve:
  - learning outcomes in MATH 220
  - failure rates and student experience
- Track proof skills through the math program

Instructors Interviews:

- Students lack "basic" (i.e. pre-requisite) logic and computational skills
  - at the start of Math 220
  - even after successfully completing Math 220

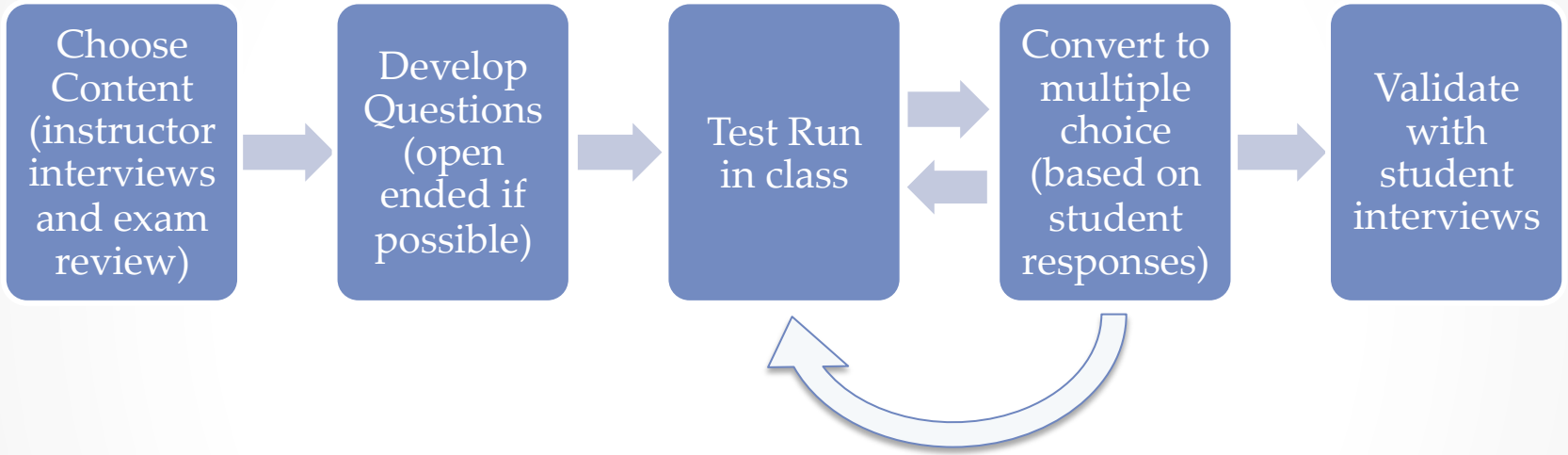
Need to assess these "basic" logic and math skills

The Basic Proof Skills Test

- Goal: Create a short (20 min), multiple-choice test to administer in Math 220 to assess skills instructors deem crucial for success in the course.
  - Focus on key observed difficulties
  - Minimize notation and technical language
  - Should correlate with performance in the course

- V1 – Sep 2010 (open-ended and multiple choice)
- V2 – Apr 2011 (open-ended and multiple choice)
- V3 – May 2011 (fully multiple choice)

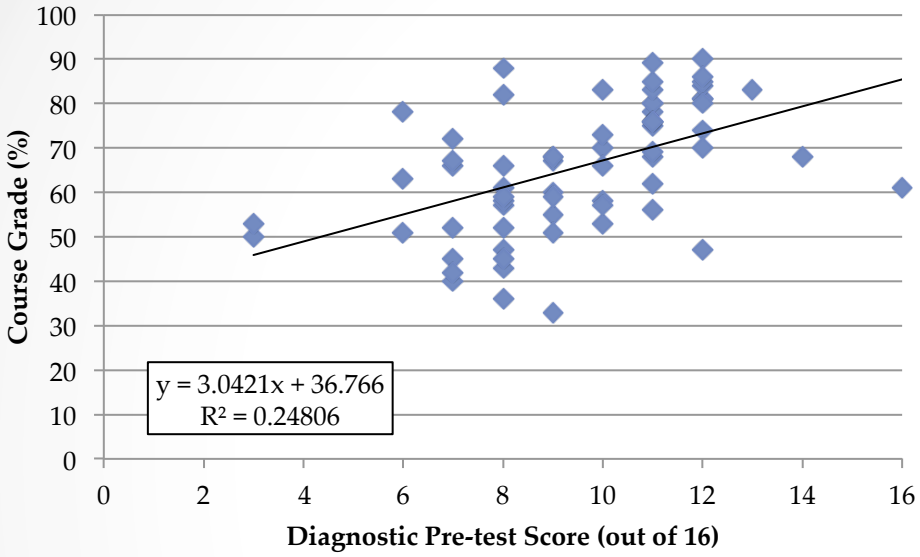
Typical Question Development

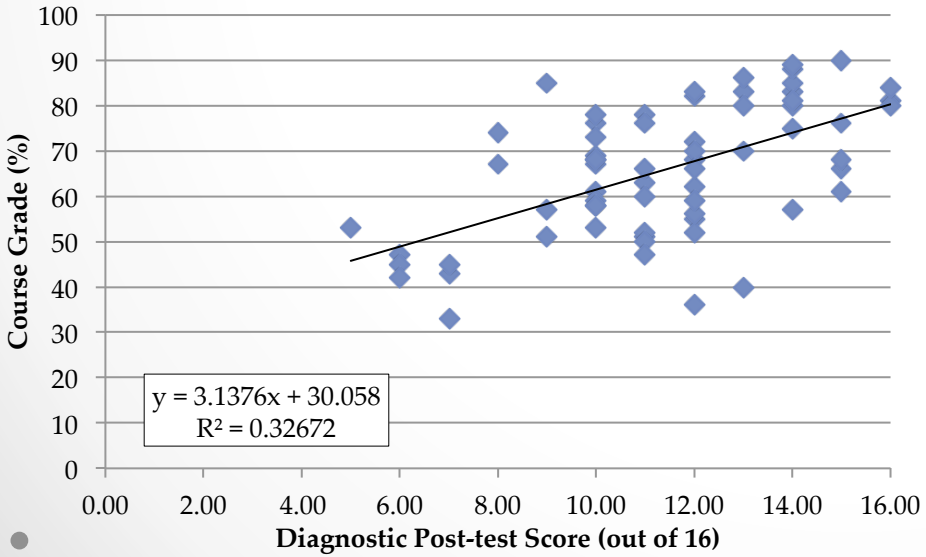


Consideration of Item Analysis Statistics and Full Test Statistics:

- Difficulty index
- Discrimination index
- Item-to-total correlation
- Item characteristic curves
- Cronbach's alpha (alpha =0.66)
- Ferguson's Delta (delta = 0.94)
- Test-retest reliability (r = 0.94)

Correlation with Grades





Pre-test and Post-test both correlate strongly with overall performance in the course

(data from 2011 Winter, Term 2)

Test Items: Relevant Algebra, Functions and Graphing

4 Questions

- Identified common errors on final exams
- Focus on absolute values and inequalities

Algebra:

Find the set of **all** values of x for which

$$|2 - x^2| < 2$$

is **true**.

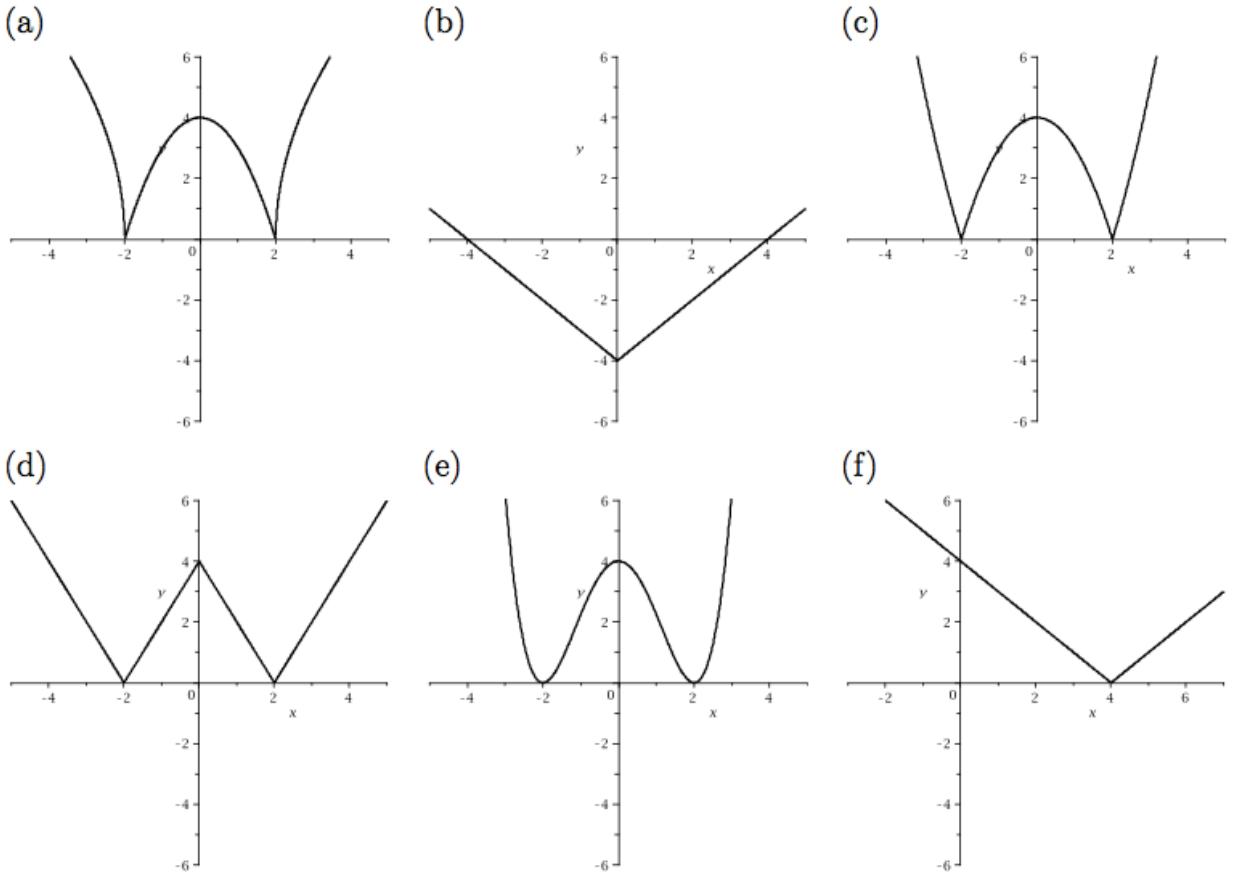
(a) $\left(0,\ \sqrt{2}\right)$ (c) $(-2,\ 0)$ (e) $\left(-\sqrt{2},\ 0\right) \cup \left(0,\ \sqrt{2}\right)$ (g) $(-2,\ 2)$

(b) $(0,\ 2)$ (d) $\left(-\sqrt{2},\ 0\right)$ (f) $(-2,\ 0) \cup (0,\ 2)$ (h) $\left(-\sqrt{2},\ \sqrt{2}\right)$

Test Items: Relevant Algebra, Functions and Graphing

Graphing:

Which of the following is a sketch of the function $g(x) = |x^2 - 4|$?



Test Items: Logic

- Logical Implication (3 questions)*
  - Equivalence to contrapositive, converse, inverse

Do the following two statements mean the same thing?

"If I am healthy, then I will come to class"

"If I come to class, then I am healthy"

(a) Yes

(b) No

- Open sentences (3 questions)

for real numbers x and y, $\sqrt{x^2 + y^2} < x$

(a) always true

(b) sometimes true

(c) never true

* Hoyles & Kuchemann (2002), Durand-Guerrier (2003)

Test Items: Logic

- Proof validation* (1 question)

Below is a statement and 3 proofs. Select the proof of the statement that is **correct and complete.**

"For any positive numbers a and b, $\frac{a+b}{2} \geq \sqrt{ab}$"

(a)

Proof: Assuming that

$$\frac{a+b}{2} \geq \sqrt{ab}$$

Multiply both sides by 2	$a+b$	$\geq$	$2\sqrt{ab}$
Squaring	$(a+b)^2$	$\geq$	$4ab$
	a^2+b^2+2ab	$\geq$	$4ab$
	a^2+b^2-2ab	$\geq$	0
	$(a-b)^2$	$\geq$	0

Which is true for positive numbers. So the assumption was true.

(Adapted from the Field-Tested Learning Assessment Guide (FLAG), Ridgway et al (2001))

* Moore (1994), Coe & Ruthven (1994), Harel & Sowder (1998), Selden & Selden (2003), Weber (2010), Mejia-Ramos & Inglis (2011), Inglis & Alcock (2012)

Test Items: Quantifiers and Definitions

- Mathematical Quantifiers* (2 questions)
  - Order of existential and universal quantifiers

True or false: There exists a real number a such that we can find a real number b such that $a - b = 4$.

(a) True

(b) False

True or false: There exists an integer x such that for every integer y, $x + y = 3$.

(a) True

(b) False

* Dubinsky (1997), Dubinsky & Tiparaki (2000), Piatek-Jimenez (2010)

Test Items: Quantifiers and Definitions

- Mathematical Definitions (3 questions)
  - Including conjunction, disjunction and negation

For a pair of integers $(a,\ b)$ we have following definition (for this test only, this is not a standard definition):

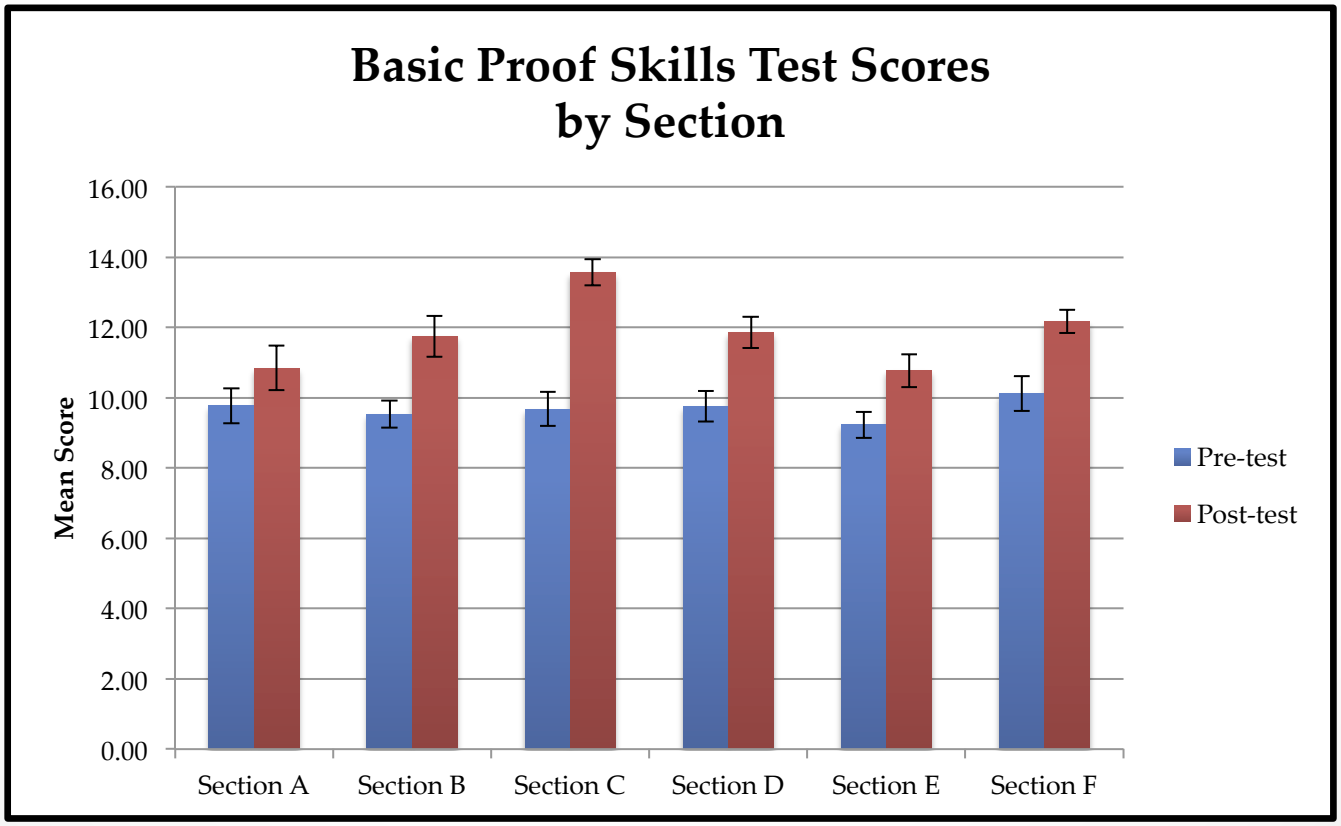
When a is even or b is odd then the pair $(a,\ b)$ is called *happy*.

Select **all** pairs below that are **happy**.

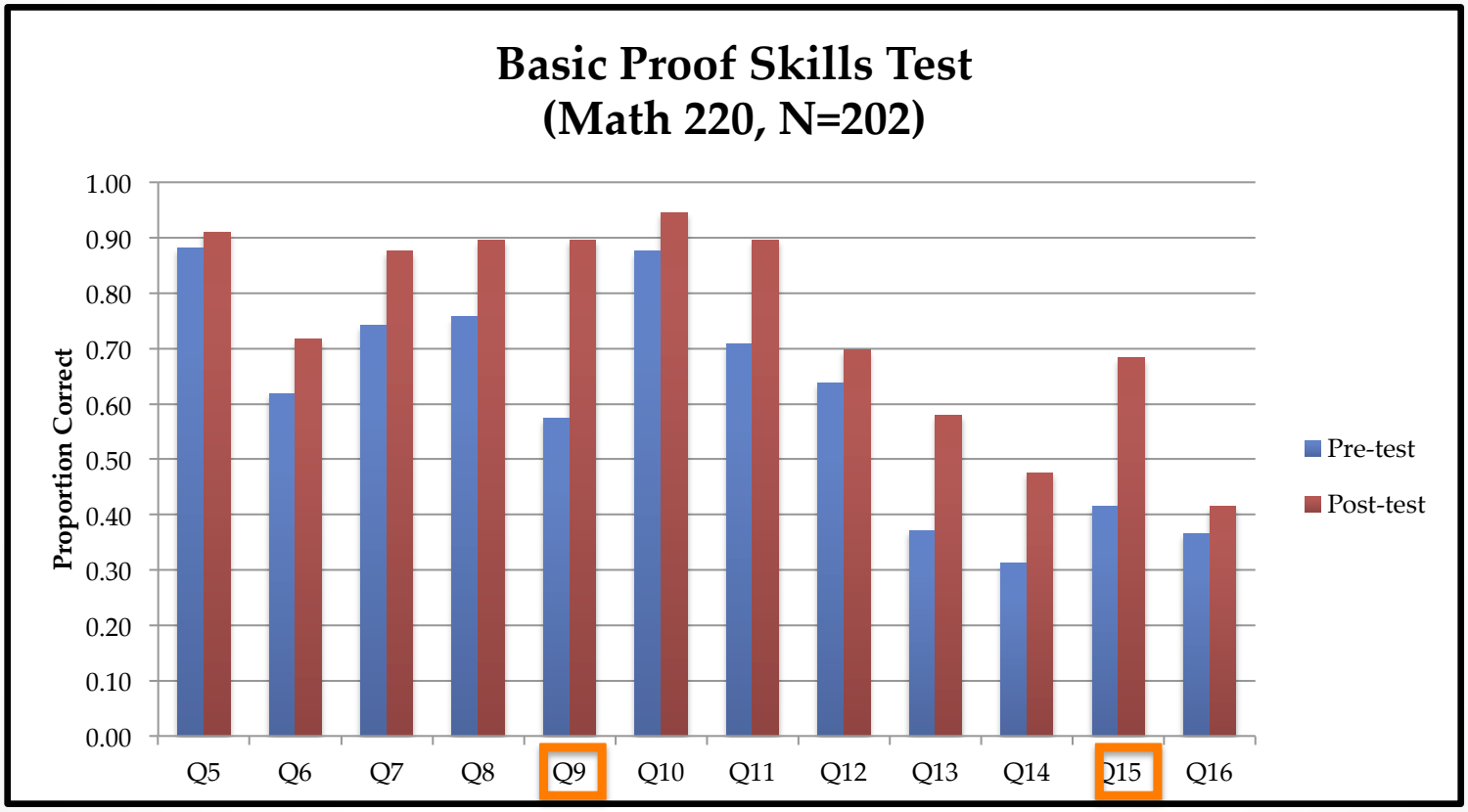
(a)	$(1,0)$	(c)	$(-2,3)$	(e)	$(3,0)$	(g)	$(-1,1)$
(b)	$(2,0)$	(d)	$(5,-1)$	(f)	$(3,-3)$	(h)	$(1,-4)$

Tracking Learning Gains

- Administered as a pre- and post-test, to track learning gains and compare instructional approaches



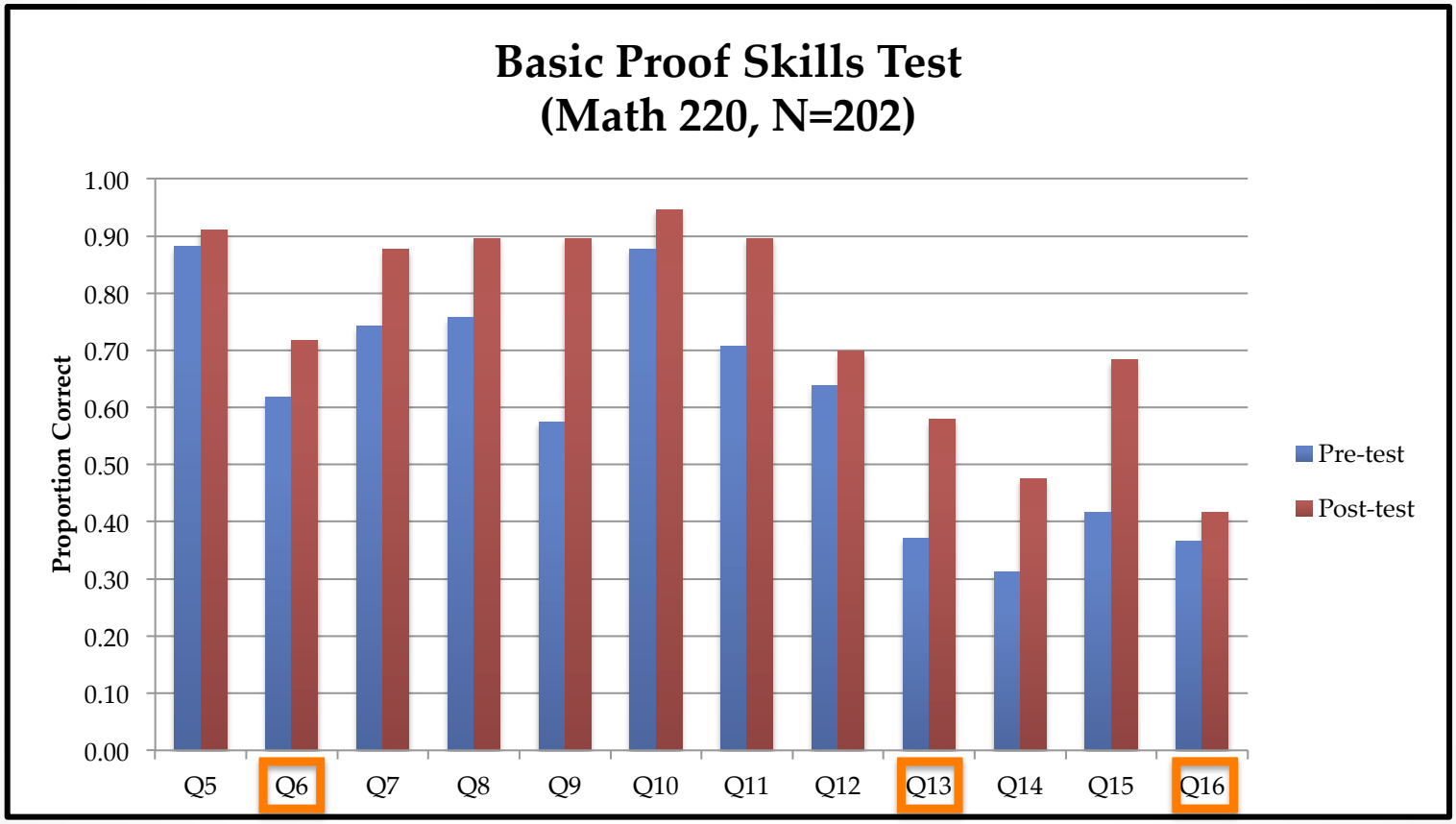
Identifying Items Responsive to Instruction



Equivalence of implication and contrapositive

Definition with disjunction

Identifying "Stubborn" Difficulties

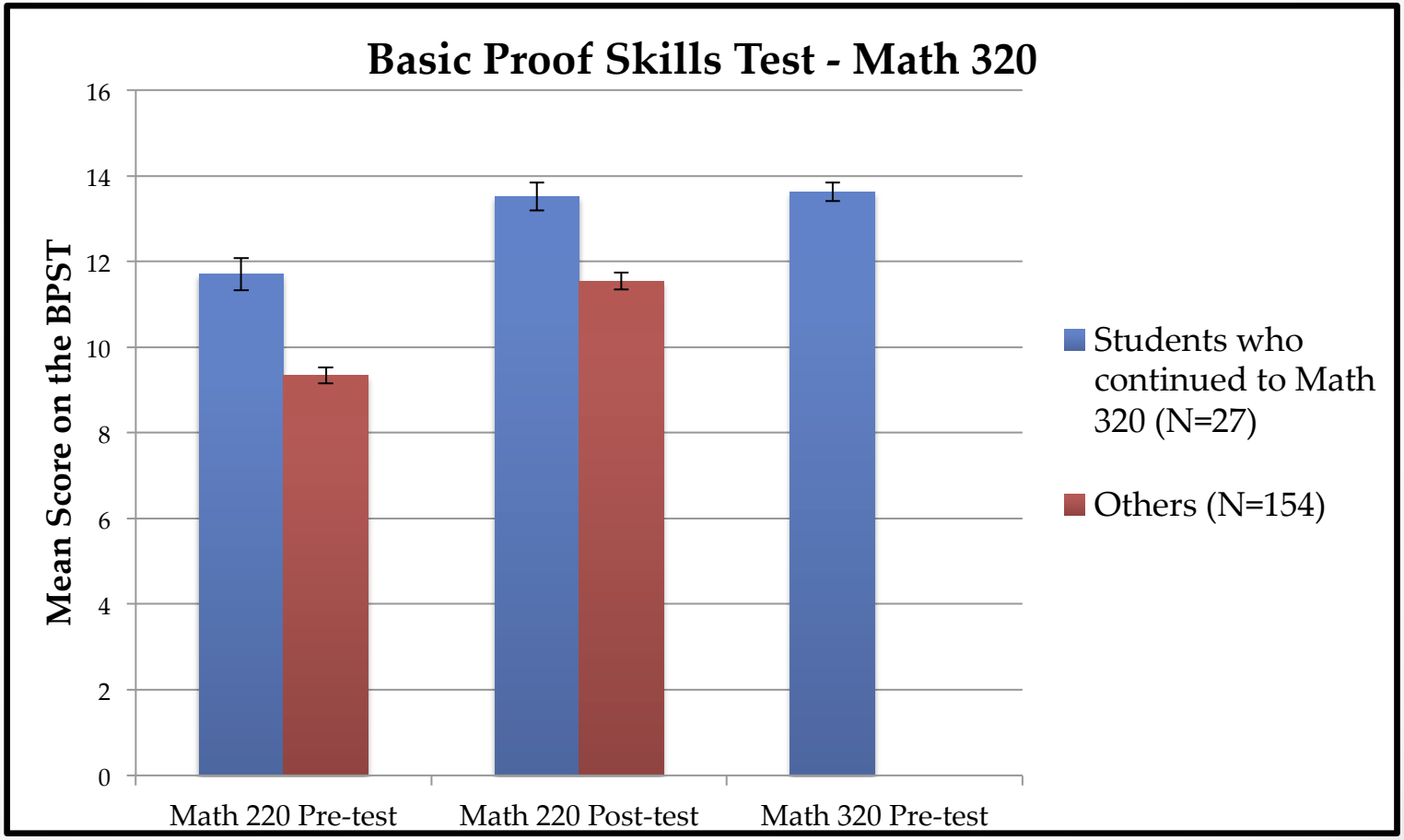


Quantifiers: There exists x such that for all y, x+y=3

Proof validation

Definition with conjunction and negation

Longitudinal Tracking and Cohort Comparison



Future Plans

- Improve the test further
  - Validity and reliability
    - Consultation with more domain experts (instructors and researchers)
    - Student validation
- Possibly extend its use to other courses or institutions
- Create a similar instrument for higher-level proof skills

References

1. Hoyles & Kuchemann (2002). Students' understandings of logical implication. Educational Studies in Mathematics, 51, 193-223.
2. Durand-Guerrier (2003). Which notion of implication in the right one? From logical considerations to a didactic perspective. Educational Studies in Mathematics, 53, 5-34.
3. Dubinsky (1997). On learning quantification. Journal of Computers in Mathematics and Science Teaching, 16, 335-362.
4. Dubinsky & Yiparaki (2000). On student understanding of AE and EA quantification. CBMS Issues in Mathematics Education, 239-289.
5. Piatek-Jimenez (2010). Students' interpretations of mathematical statements involving quantification. Mathematics Education Research Journal, 22, 41-56.
6. Ridgway et al (2001). Assessing mathematical thinking via FLAG. In The Teaching and Learning of Mathematics at University Level, 423-430.
7. Coe & Ruthven (1994). Proof practices and constructs of advanced mathematics students. British Educational Research Journal, 20, 41-53.
8. Moore (1994). Making the transition to formal proof. Educational Studies in Mathematics, 27, 249-266.
9. Harel & Sowder (1998). Students' proof schemes: Results from exploratory studies. CBMS Issues in Mathematics Education, 7, 234-283.
10. Selden & Selden (2003). Validations of proofs written as texts: Can undergraduates tell whether an argument proves a theorem? Journal for Research in Mathematics Education, 34, 4-36.
11. Weber (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, 12, 306-336.
12. Mejia-Ramos & Inglis (2011). Semantic contamination and mathematical proof: Can a non-proof prove? Journal of Mathematical Behavior, 30, 19-29.
13. Inglis & Alcock (2012). Expert and novice approaches to reading mathematical proofs. Journal for Research in Mathematics Education, 43, 358-390.

Question	Pre-test			Post-test		
	Difficulty Index	Discrimination Index	Item-to-total Correlation	Difficulty Index	Discrimination Index	Item-to-total Correlation
1	0.62	0.34	0.06	0.72	0.48	0.26
2	0.61	0.55	0.25	0.68	0.60	0.30
3	0.64	0.62	0.24	0.71	0.48	0.24
4	0.31	0.48	0.17	0.48	0.69	0.32
5	0.88	0.14	0.03	0.91	0.23	0.20
6	0.62	0.52	0.16	0.72	0.31	0.12
7	0.74	0.44	0.23	0.88	0.33	0.33
8	0.76	0.31	0.14	0.90	0.38	0.36
9	0.57	0.49	0.16	0.90	0.25	0.29
10	0.88	0.26	0.19	0.95	0.10	0.13
11	0.71	0.52	0.30	0.90	0.30	0.33
12	0.64	0.44	0.11	0.70	0.58	0.38
13	0.37	0.53	0.14	0.58	0.51	0.23
14	0.31	0.23	0.07	0.48	0.40	0.12
15	0.42	0.46	0.10	0.68	0.63	0.28
16	0.37	0.27	0.01	0.42	0.70	0.31

Cronbach's Alpha

Pre-test: 0.48

Post-test: 0.66

Ferguson's Delta

Pre-test: 0.93

Post-test: 0.94

Average (corrected) Point-biserial correlation:

Pre-test: 0.15

Post-test: 0.26

Test-retest Reliability (computed correlation of item difficulty indices for two separate term pre-tests)

Correlation coefficient: 0.944