Assessing basic skills for mathematical proof

Sandi Merchant & Andrew Rechnitzer

CWSEI End-of-Year Event
Apr 29, 2011

Abstract

MATH 220: "Mathematical Proof" is a key component of our Math curriculum, but failure rates are high and both students and instructors have expressed dissatisfaction with learning. We are developing a basic proof skills diagnostic test to assess the skills of incoming students as well as to track the learning of these key skills over the term. Here, we present the development of the test so far as well as some of the more interesting results in terms of student difficulties and learning as well as the predictive power of the test.

We have found that there are substantial learning gains over the course in some of these skills, but that some difficulties persist and may require new materials and methods to address. In terms of performance in the course, we see fairly high predictive power for such a short test (20 min.). In addition, the test appears to measure skills that differ from those measured in a our calculus II courses.

Background

- Math 220: "Mathematical Proof"

  "Sets and functions; induction; cardinality; properties of the real numbers; sequences, series, and limits. Logic, structure, style, and clarity of proofs emphasized throughout."

- "Gateway" course for study in pure math
  - Required for most non-honours math and nearly all statistics majors
  - Pre-requisite for courses in: Algebra, Analysis, Geometry and Number Theory

- High failure rates ($> 25\%$) and dissatisfaction by instructors of Math 220 and subsequent courses.

- MATH 220 was selected for transformation under CWSEI, starting May 2010.

The "Basic Proof Skills Diagnostic Test"

The goal: to create a diagnostic tool that

- Can be administered at the start of term to:
  - Assess incoming students' basic knowledge and skills, to guide instruction.
  - Predict student performance in MATH 220 (identify "at risk" students).
- Can be administered again at the end of term to:
  - Track learning over the course and the effects of course transformation elements.
- Is short and easy to score.
  - Less than 15 multiple choice questions, 20 min. test duration.

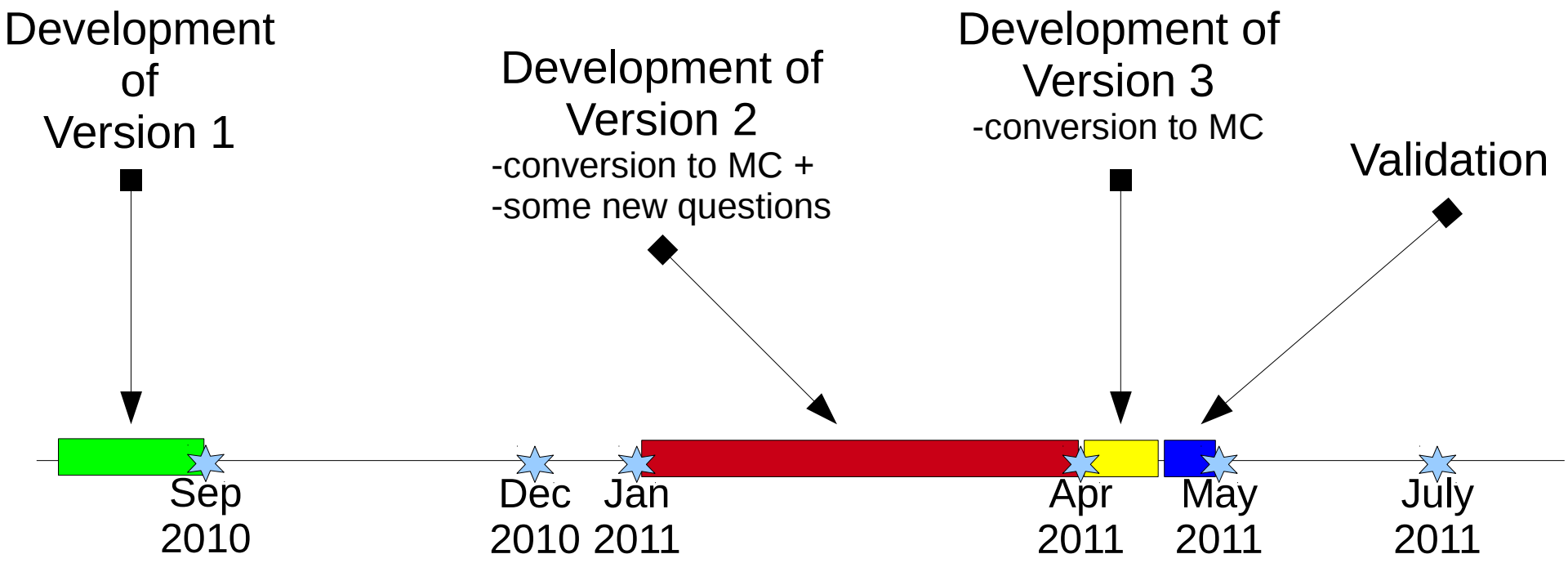
Development of the Diagnostic

- We identified key basic skills based on:
  - course resources: textbook, lecture notes, homework
  - interviews with instructors (MATH 220 and subsequent courses)
  - past final and midterm exams
  - observations during problem-solving sessions

- Three main types of key skills were identified:

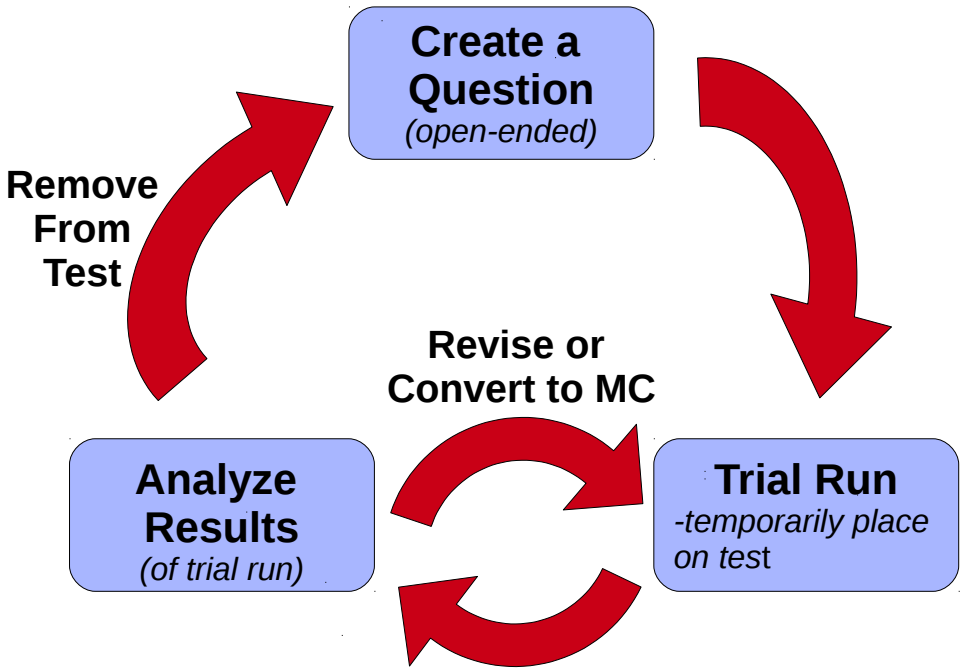
  (i) logic and reasoning
  (ii) relevant algebra and graphing
  (iii) mathematics reading and comprehension

Timeline



= trial run in class

Question Development



Questions were analyzed and accepted/rejected based on:

- clarity or ambiguity
- quality and variety of incorrect responses
- distribution of responses (ceiling effects, etc.)
- correlation with performance on final exam

Examples from the Proof Diagnostic Test
(i) logic and reasoning

Basic logic:

- Consider the statement

  "If it is difficult, then it makes me stronger"

  Select all statements below that **mean the same thing** as the above statement:

  (a) If it is not difficult, then it does not make me stronger.
  (b) If it makes me stronger, then it is difficult.
  (c) If it makes me stronger, then it is not difficult.
  (d) If it does not make me stronger, then it is difficult.
  (e) If it does not make me stronger, then it is not difficult.

  - Sept 2010: 44% of students incorrectly circled (a)

Examples from the Proof Diagnostic Test
(i) logic and reasoning

Quantitative reasoning skills:

- Indicate whether the statement

$$\sqrt{x^2 + y^2} < x$$

is:

(a) always true: true for any choice of variable or object
(b) sometimes true: true for some variable or object choices, but not for all choices, or
(c) never true: not true for any choices of the variable or object.

  - Sept 2010: only 48% of students correctly answered (c).

Examples from the Proof Diagnostic Test
(i) logic and reasoning

Identifying correct vs. incorrect proofs:

- Choose the BEST proof of the statement

$$(a+b)/2 \geq \sqrt{ab}$$

  - 2 of 3 options were incorrect proofs
  - Sept 2010: only 33% answered correctly
  - 56% chose a "proof" that began by explicitly assuming the result to be proved).

Examples from the Proof Diagnostic Test
(ii) relevant algebra and graphing

Algebra:

- Find all values of k such that

$$(k-2)^2 - (k-1)^2 - 6k < 5k - 3$$

is true.

  - Sept 2010: only 49% correctly solved.

Graphing:

- Draw a rough sketch of the function

$$g(x) = |x^2 - 4|$$

on the axes below.

  - Sept 2010: only 56% drew correct graphs.

Examples from the Proof Diagnostic Test
(iii) math reading and comprehension

Quantifiers (in plain English):

- Select all statements below that are **true**:

  (a) There exists an integer x such that we can find an integer y such that $x + y = 3$.

  (b) There exists an integer x such that for every integer y, $x + y = 3$.

  (c) For every integer x, we can find an integer y such that $x + y = 3$.

  (d) For every integer x, and for every integer y, $x + y = 3$.

  - Sept 2010: 35% of students incorrectly circled (b).

Mean Calculus II grade is 78%

Pre/Post Results

Comparison of end-of-term (post-test) scores vs. start-of-term (pretest) scores found:

Significant improvement in some skills

- T/F: "If it does not make me stronger then it is not difficult" means the same thing as "If it is difficult, then it makes me stronger."

$$\boxed{75\% \rightarrow 99\%}$$

- Always/sometimes/never true: $(x+y)^2 = x^2 + y^2$.

$$\boxed{63\% \rightarrow 87\%}$$

- Factor 119 into products of prime numbers.

$$\boxed{24\% \rightarrow 63\%}$$

Pre/Post Results

Some difficulties remain at the end of term

- T/F: There exists an integer x such that for every integer y, $x + y = 3$.

$$\boxed{65\% \rightarrow 71\%}$$

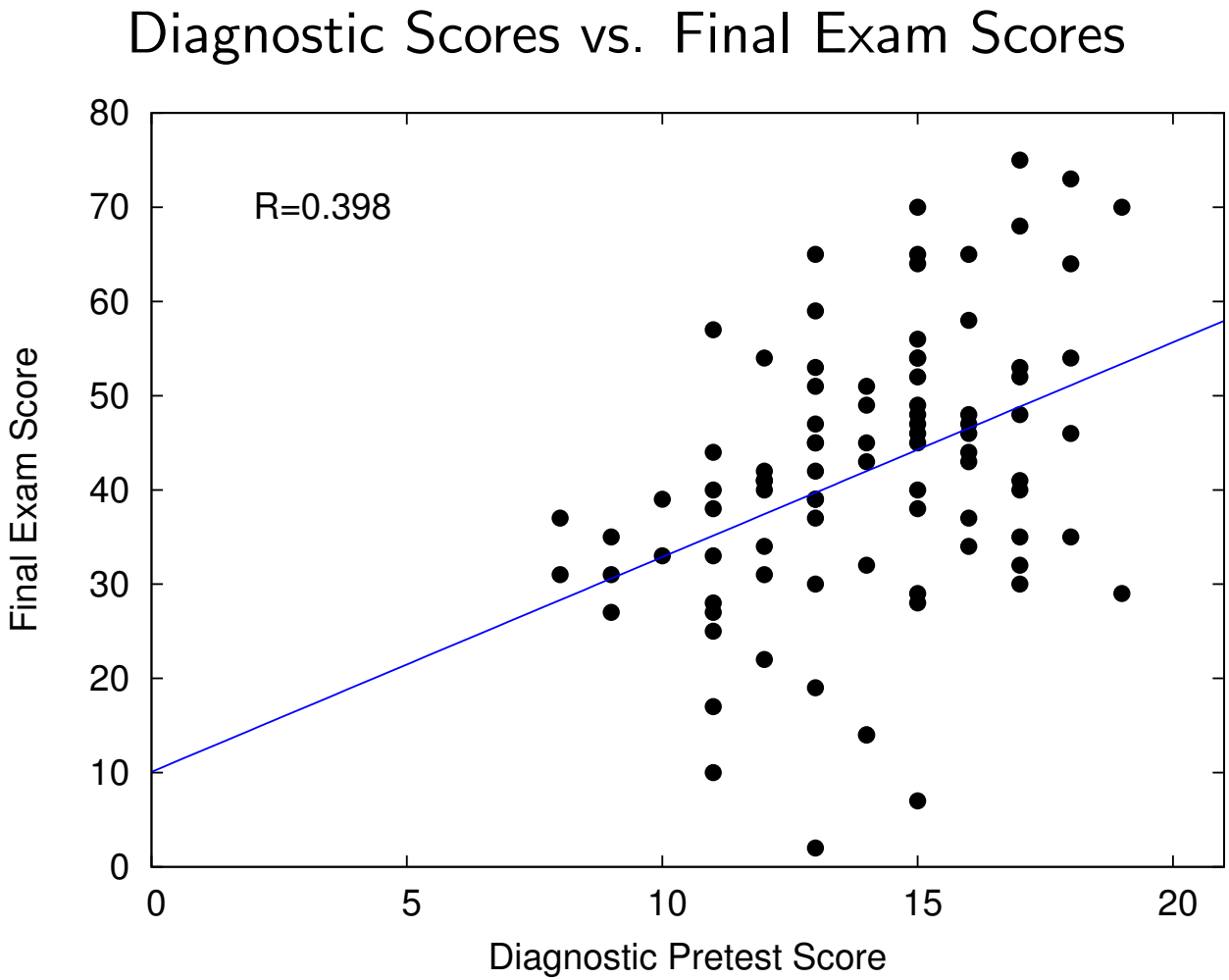
- Always/sometimes/never true: $\sqrt{x^2 + y^2} < x$.

$$\boxed{48\% \rightarrow 67\%}$$

- Choose the best proof of "For any positive numbers a and b, $(a + b)/2 \geq \sqrt{ab}$."

$$\boxed{33\% \rightarrow 51\%}$$

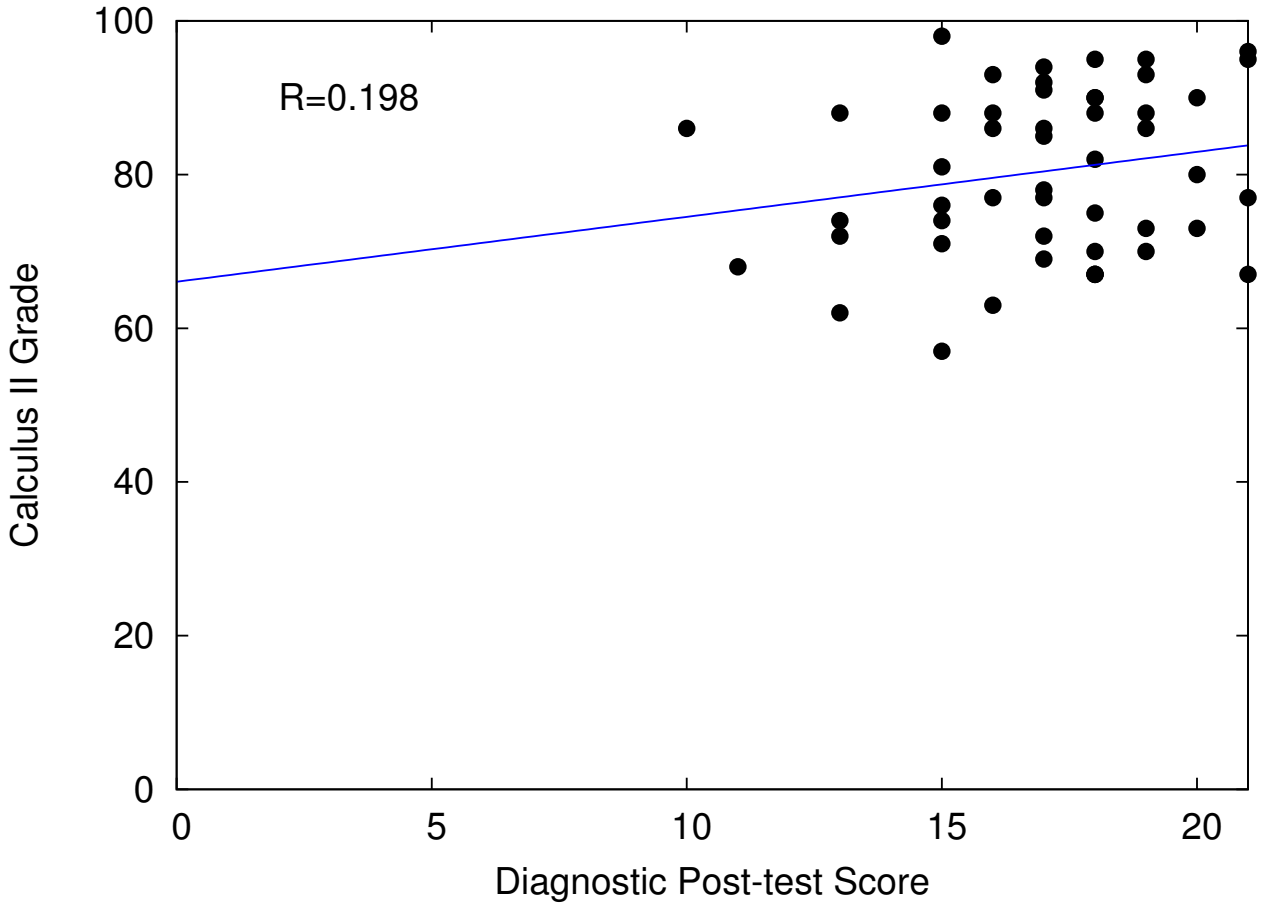
Does the test accurately predict success?



*Fall 2010, both sections pooled.

*Diagnostic scores are from the start of term diagnostic test.

Does this test tell us anything different from their Calculus II grade?



*Correlation between diagnostic scores and Calc II grades is low.

*Diagnostic test measures skills that differ from Calc II.

Future Work

- Continue to improve the test (analyze, add and remove questions).
- Validate questions with student interviews.
  - Are questions being interpreted they way we expect?
  - Are correct answers being chosen for valid reasons?
  - What is the misconception that causes each incorrect answer?
- Develop a more advanced proof skills test to study retention to later courses.
- Develop course materials and activities to address deficiencies identified by the test.