Motivation and Goals

- Integrate cognitive and developmental research with psychometric advances, to produce a technically sound assessment system focused on the impact of program on student knowledge and performance (Pellegrino, Baxter, Glaser, 1999; Shepard, 2000).

- Explicit descriptions of student understanding can be used to
  - Monitor progress toward targeted learning goals,
  - Provide visible examples of developing competence, and explicated gaps in understanding;
  - Suggest activities for promoting further development.

- Our project focuses on several crosscutting areas in mathematics. The first, pilot, area is proportional reasoning.
Why Proportional Reasoning?

- Rich research resources (e.g. Kaput & West, 1994; Noelting, 1980ab; Lamon, 1999ab; Lesh, Post & Behr, 1988).
- Not dependent on any one curriculum; appears in all “exemplary” mathematics curricula (US Dept of Ed, 1999).
- Prominent in Math Standards (NTCM, 1989, 2000), and an important precursor to algebraic thinking.
- Prominent in Science Standards (NRC, 1999) and central to scientific literacy and understanding (AAAS, 1993).

Some Antecedents

Our work is related to

- Past thinking about performance assessment in math and science (e.g. Carpenter and Fenema, 1999; Case, 1996; Hunt and Minstrell, 1999).
- Curriculum-embedded assessment projects like ACER’s “Progress Maps” (Masters & Forster, 1996); and the BEAR project (Wilson & Sloane, 2000).

But we want to be informative about the impact of program on student knowledge and performance

- US curriculum heterogeneity argues against simply ordering tasks and students along a single linear scale.
- Instead, identify stable clusters of strategies, representations and understandings characteristic of various levels of development.
Design Considerations

- A cognitive-developmental map of Proportional Reasoning
  - Working definition
  - List of typical problem types and contexts
  - Description of aspects of student performance that constitute evidence of understanding
  - Description of a developmental sequence as students’ expertise progresses
- A set of tasks to elicit knowledge-based performance variation
- Task scoring criteria to emphasize relevant individual differences
- A summary of performance so teachers can see how students’ performances are related to the cognitive developmental map within and across grade levels.

Working Definition of Proportional Reasoning

- …A Psychological Construct: “focuses on describing, predicting or evaluating the relationship between two relationships…” (Piaget and Inholder, 1975, as cited by Lesh, Post and Behr, 1992).
- …With Mathematical Content: Reasoning about two related mathematical models:
  - Equal ratios: \( \frac{a}{b} = \frac{c}{d} \)
  - Linear scaling: \( y = k \cdot x \)
- Recognizing situations in which a proportional or multiplicative relationship is, or is not, relevant. (i.e. real-world modeling)
- Understanding that both component measures in a ratio are varied but in such a way that the original ratio relationship remains invariant. (i.e. covariance and invariance)

Most curricula (!) and research studies (!!) do not succinctly define it!
Typical Problem Types and Contexts

- Computational proportional reasoning: Vergnaud’s (1983) “isomorphisms of measures”:

<table>
<thead>
<tr>
<th>Measure 1</th>
<th>Measure 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case 1</td>
<td>A</td>
</tr>
<tr>
<td>Case 2</td>
<td>C</td>
</tr>
</tbody>
</table>

Two problem types:
- Missing value: given 3 terms (e.g. A,B,C), find the 4th (e.g. D)
  A stack of four pennies is 5mm high. How high would a stack of ten pennies be?
- Comparison: A:C vs. B:D, or A:B vs. C:D.
  My table has seven people and 2 12” pizzas. Your table has twelve people and 3 12” pizzas. If each table shares fairly, would a person at my table or your table get more pizza?

Observable Performance Variation: Strategy

Anne and Kathy each bought the same kind of bubble gum at the same store. Ann bought two pieces of gum for six cents. If Kathy bought eight pieces of gum, how much did she pay?

<table>
<thead>
<tr>
<th>Bubble Gum (pc’s)</th>
<th>Cost (cents)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anne</td>
<td>2</td>
</tr>
<tr>
<td>Kathy</td>
<td>8</td>
</tr>
</tbody>
</table>

Rule of three algorithm: 2 × ?? = 6 × 8.

Between (functional): Calculate unit rate 6 ÷ 2 = 3 from one case and apply to the other: 8 × 3 = 24.

Within (scalar): Calculate factor-of-change 8 ÷ 2 = 4 in one measure space and apply to the other: 6 × 4 = 24.

Coordinated build-up (scalar): (2, 6) + (2, 6) + (2, 6) + (2, 6) = (8, 24):
  number of terms determined by trial-and-error.

Non-multiplicative: 6 – 2 = 4; ?? – 8 = 4.
Accessing Conceptual Understanding (e.g. Lamon, 1999)

- **Absolute vs. relative comparison**: Store A advertises a sale of $10.00 off. Store B advertises a sale of 10% off. If both stores sell the same things, which store offers the best sale?

- **Recognizing presence/absence of covariation**: John and his brother drove to school in 10 minutes. How long does it take John to drive there by himself?

- **Understanding covariance/invariance**: Each table in Mr. Trent’s science class seats 3 children. He places 2 candy bars on each table for the students to split fairly. Is it better to get your fair share from your own table, or push all the tables together and share all the candy fairly among all the children?

- **Understanding real-world contexts**: Suppose you have 5 cookies and three children. If they share fairly, how many cookies does each child get? What if they are sharing 5 books instead of 5 cookies?

- **Part-whole vs part-part**: In a group consisting of 24 boys and 6 girls, what percentage of the class are girls?

---

**Observational and Developmental Challenges**

- Strategies similar across problem type and domain.

- Strategy use inconsistent across tasks and sometimes within tasks.

- “Within” easier than, and preferred to. “between”, even when numbers work out better for “between”.

- Students (and adults) often view as quite different:
  - \( \frac{a}{b} = \frac{c}{d} \): sharing, grouping, rate comparison and mixing; vs.
  - \( y = kx \): shrinking/stretching, familiar intensive rates (e.g. mph)

- Many factors influence strategy choice: nice numbers, location of unknown, discrete vs. continuous.

- When the numbers are not “nice” then \( \frac{a}{b} = \frac{c}{d} \) or \( y = kx \) may be tried, or else fall back on additive model.

- **Strategy changes to “repair” unanticipated difficulties…**
**Major Stages of Development**

- **Qualitative:** Enough knowledge about quantity to answer questions about more/less (e.g. which is sweeter?) and fairness (e.g. sharing pizza).
- **Early Quantitative:** Early attempts often involve constant additive differences \((a - b = c - d)\) and counting up and down.
- **Recognize Multiplicative:** Recognize differences change with size of numbers; no explicit covariance/invariance model. Rely on replicating, build-up, and unit-factor strategies.
- **Accommodate Covariance / Invariance:** Explicit covariance/invariance model. Still rely on coordinated build up, unit factor, and other scalar approaches; when strategies fail to generalize, may fall back on additive differences.
- **Scalar and Functional:** Understand invariant nature of relationships between changing quantities. General model for proportional reasoning allowing choice among strategy repertoire more or less independently of semantics, context, or cover story.

_No definitive studies of development, grades 5–8, exist!_

**Statistical Modeling and Scoring**

- The pattern of evidence _within_ a task ⇒ inferences about strategies, representations and understandings for that task.
- The pattern of strategies, etc., _across_ tasks ⇒ inferences about developmental stages.
- The developmental stages suggest a “fuzzy sets” or “latent class” approach (e.g. Moore, Dixon and Haines, 1991).
- However, the latent classes can be left unordered and the “ordering”, if there is any, can emerge from the fitted model.
- Right/wrong and partial credit _not_ as important as identifying, e.g. strategy choice/implementation.
- Variations on IRT/latent class models allow scoring for strategy rather than right/wrong (e.g. Rijkes, 1996). Mathematically similar to recent multiple-rater models (Patz, Junker, and Johnson, 2000).
Conclusions

- **The challenge:** incorporating what we know about student performance variation and its relationship to student understanding in ways that can guide assessment design, scoring and reporting.

- We have begun to meet this challenge in proportional reasoning, by providing
  - A *definition* of proportional reasoning;
  - Some *problem types and contexts* in which they are situated;
  - Some *observable features of student performance that are indicative of emerging understanding*, based on our synthesis of the literature;
  - A set of *developmental stages* through which we expect students to pass as their understanding grows.

- This provides a set of expectations (or hypotheses) that can guide task design and empirical evaluation in the next phase of our work.