AI Generated Executive Summary of the Publication Visualizing and Reporting Content-Referenced Growth on a Learning Progression

The publication, "Visualizing and Reporting Content-Referenced Growth on a Learning Progression," by Briggs et al. (2025), introduces and illustrates a novel framework for Content-Referenced Growth (CRG) score reporting in educational assessments. This approach aims to provide teachers, parents, and students with more instructionally meaningful insights into student learning and growth by connecting quantitative test scores to qualitative stages along a research-based learning progression.

I. Core Problem: Limitations of Current Score Reporting

The authors identify a significant challenge in educational measurement: the effective presentation of test results. They argue that traditional score reports, often characterized as "purely normative" or "criterion-referenced" (e.g., "growth to standard"), fall short in providing actionable insights for instruction.

- Lack of Substantive Insights: Numeric growth targets (e.g., percentile ranks, scale score points) do not convey *what* a student has learned. "Most existing score reports from interim assessment vendors are designed to contextualize, at a high level, what it is that students are likely to know and be able to do given their most recent scale score. The focus is not placed on how this has changed across occasions and how this change can be interpreted."
- Focus on Monitoring vs. Formative Use: A reliance on numerical targets can lead to assessments being perceived primarily as tools for "monitoring and evaluating: a student has either met or not met a given target." This summative focus undermines the potential for formative use of assessment data.
- Ambiguous Meaning of Scale Scores: Scale score units often lack intuitive meaning, making it difficult for teachers and parents to understand the qualitative implications of a score increase.

II. The Content-Referenced Growth (CRG) Approach

The central premise of the article is that "when an educational assessment has been designed to support teachers in using test results for formative purposes, then (1) score reporting should emphasize growth as much (or more) than it does status, and (2) this should be done in a way that encourages teachers to connect their interpretations of student growth to the content of the assessment."

- Qualitative Interpretation of Growth: Instead of merely stating a student has "grown 60 points in math," a CRG interface would "help a teacher to convey that a student has gone from solving problems that involve a part-whole conceptualization of fractions, to solving problems that require the student to locate fractions on a number line." This emphasizes the conceptual development rather than just a numerical change.
- Facilitating Conceptual Adjustments: The CRG approach aims to move teachers beyond "procedural adjustments" (e.g., creating student groups) to "conceptual adjustments" in

their instruction, which are "most likely to help students learn." This is achieved by focusing "teacher attention on the developmental path that students take as they learn a big picture topic."
- Theory of Action: The CRG framework operates on a high-level theory of action (Figure 1), where improved student learning outcomes are achieved as teachers adjust their instruction based on inferences from the CRG reporting interface. This also seeks to contribute to teachers' professional learning and improve their attitudes toward the usefulness of assessment.

III. Four Core Elements of the CRG Framework

The CRG approach integrates four key elements to provide instructionally meaningful interpretations of growth:

1. Item Mapping: This involves using qualitative distinctions among assessment items to make sense of locations along a score scale. Its "more important purpose...is to provide a qualitative interpretation of the numerical distances between scale score locations with respect to the qualitative differences of exemplar items at these locations." This helps teachers understand *what* a score difference means in terms of content.
2. Learning Progressions (LPs): These are "descriptions of the successively more sophisticated ways of thinking about a topic that can follow one another as children learn about and investigate a topic over a broad span of time." LPs provide a theoretical basis for ordering item difficulty and offer "qualitatively distinct levels or 'waypoints'" that represent significant steps in understanding.
3. Targeted Test Design: This element emphasizes designing tests with "overlapping content across temporal occasions" (e.g., grades) rather than entirely unique content. This allows for disentangling differences in student ability from item difficulty and facilitates the creation of items that discriminate between LP levels.
4. Vertical Scale Calibration (using Rasch Model): The Rasch Model is preferred for calibrating tests on a common vertical scale because it allows for "invariance of comparisons among persons to the choice of items used to define a reference interval." This means the interpretation of distances along the scale does not depend on the specific items used.

IV. Illustrative Example: Understanding Fractions Learning Progression

The paper illustrates the CRG framework using a prototype for an "Understanding Fractions Learning Progression" based on Kieren's five conceptualizations of fractions. This LP has four levels:

- Level 1: Part-Whole: Understanding fractions as a specified number of parts out of a total, often visually.
- Level 2: Quotient (Fair Shares): Understanding fractions as equal partitions or division expressions, and the iteration of unit fractions.

- Level 3: Measurement (Number Line): Understanding fractions as unique numerical values that can be placed on a number line, allowing for ordering and understanding equivalent fractions for addition/subtraction.
- Level 4: Operator (Multiply and Divide): Interpreting fractions as operators that produce a new proportional value through multiplication and division.

Empirical Validity Evidence:

- Curriculum Alignment: There is a "moderate to strong association" (rank correlation of .68) between the chronological order of fraction-related lessons in Curriculum Associates' i-Ready Classroom Mathematics and the LP level of assessed skills, suggesting that "as students move up in grade levels they are indeed more likely to receive instruction about fractions that would involve the more sophisticated conceptualization of fractions found in the higher levels of the LP."
- Item Difficulty Alignment: "As the LP level increases (from left to right on the horizontal axis), so does the average item difficulty. The rank correlation is .60." This empirically supports the hypothesized ordering of the LP levels based on item difficulty on the i-Ready Diagnostic. Even controlling for grade level, items coded to higher LP levels were significantly harder.

V. CRG Reporting Prototype Features (i-Ready Diagnostic Context)

The prototype is built using data from the i-Ready Diagnostic, a K-12 adaptive assessment administered multiple times a year. Key features include:

- Dynamic and Interactive Interface: Teachers can toggle between student and class-level views, order students, add color gradients, and zoom into level-specific LP interpretations with exemplar items.
- Visual Connection: The interface visually connects i-Ready Diagnostic scale scores (left axis) with the four LP levels (right axis) using horizontal lines and optional color coding.
- Qualitative Descriptions and Exemplar Items: Clicking on an LP level provides "a bulleted summary of what a student would and would not be expected to understand" and an "exemplar item they would be expected to be able to solve correctly." The formal LP names are rephrased for teachers (e.g., "Fair Shares" for "Quotient," "Number line" for "Measurement," "Multiply and Divide" for "Operator").
- Focus on Growth Trajectories: In the student view, the prototype maps an empirical growth trajectory across test occasions, emphasizing how a student moves through the LP levels over time.

VI. Pilot Test Findings with Teachers

A small-scale pilot test with seven practicing teachers (grades 3-5, i-Ready Diagnostic experience) using a "think-aloud" protocol yielded positive results.

- Meaningful Interpretations: Teachers "were able to use the prototype to make connections between i-Ready scale scores and LP levels to support inferences about student growth." While some teachers initially defaulted to prior numerical growth interpretations, others directly made content-referenced interpretations, for instance, noting a student had "a decent understanding of how it [fractions] can be represented on a number line. They are seeing the relationships between the fractions and decimals…they're not really ready for the multiplication and division."

- LP Understanding: All teachers could follow the hierarchical distinctions in the LP levels. However, some suggested adjustments to the "grain size" of levels or the order of middle levels, particularly noting that the "Fair Shares" (Level 2) exemplar item seemed more cognitively demanding than the "Number Line" (Level 3) exemplar, which aligns with how these concepts are introduced in CCSS-M. This highlights that "the levels of the fractions LP were not designed to be identical to the grade level ordering of standards found in the CCSS-M."
- Envisioned Use Cases: Teachers identified three main ways they could use the prototype:

1. Grouping Students: For targeted instructional activities.
2. Parent Communication: To easily explain student progress in meaningful terms during conferences.
3. Teacher Professional Learning: To deepen content knowledge, especially for newer teachers or those less familiar with math standards.

VII. Strengthening Validity and Future Directions

The authors acknowledge areas for improvement and future research to strengthen the validity of CRG interpretations:

- Integration with Professional Development: Effective implementation would require "initial and ongoing professional development and collaboration" to help teachers shift from a "count up points" framing to a learning progression approach.
- Choosing Exemplar Items: Care is needed to select exemplar items that truly represent LP levels and avoid misinterpretations (e.g., a single item becoming the sole instructional target). The authors suggest choosing items by a "standard criterion" and offering "a 'zoomed in' view of each level that includes an additional three items that span most of the range of scale scores."
- Item Design: While current i-Ready items are mostly selected-response, an LP design approach "may require assessment tasks that look quite different than traditional selected-response items" (e.g., ordered-multiple choice, AI-enabled chatbots for eliciting reasoning).
- Triangulation: CRG results should be "a starting point for inquiry rather than a final determination" due to measurement error and the probabilistic nature of the LP-item difficulty relationship. Teachers should use "triangulating evidence from tasks that do not have the same constraints as the items found on a standardized assessment."
- Scale Robustness: Further research is needed to ensure that vertical scales built on a broad domain (like overall math) can be robustly interpreted for specific sub-domains (like fractions).
- Empirical Research: Future studies should explore how teachers interact with real student data, whether the interface leads to desired inferences and instructional steps, and the impact of CRG on teacher attitudes and student outcomes.

VIII. Conclusion

The CRG approach represents a significant advancement in score reporting for large-scale assessments, moving towards "interactive, not static" reports that provide "meaningful evidence about student growth." It aims to empower teachers with qualitative, content-referenced interpretations of student progress, fostering deeper understanding of learning and supporting more effective instructional adjustments. The authors emphasize that the goal is not to provide definitive answers, but to prompt "substantive questions about student learning," thereby breaking from the status quo of less actionable score reporting.