



Building a Next Generation Diagnostic Assessment and Reporting System within a Learning Trajectory-Based Mathematics Learning Map for Grades 6-8
Jere Confrey, Principal Investigator

One of the greatest challenges in middle grades mathematics education is serving the diverse preparation and needs of students. Our proposal addresses the assessment strand as its primary focus and seeks to secure funding to conduct late stage design and development work on a next generation diagnostic assessment system for middle grades mathematics. It is grounded in research on learning trajectories around nine big ideas, and designed to inform instructional practices within competency-based or personalization-based models of school change (as well as other more traditional approaches). We introduce these school change models in some detail, to assist reviewers in recognizing the urgency of the need for a measure of precise learning objectives to guide system-wide reform decisions.

Establishing a National Need

Within a single classroom, students' preparation can range from 1-2 grades or more below grade level to 1-2 or more grades above (Carlson et al., 2011) and with SES accounting for 15% of the variation in scores (OECD, 2012). Schools are wary of tracking, for its well-documented dampening effects on achieving equity for low performers (Burris, 2014; Mehan, 2015; Oakes, 1990, 2015). As a result, schools have turned to differentiated instruction. There is widespread support for this concept, but little solid research that articulates proven instructional strategies, including the successful use of data (Delisle, 2015). Digital platforms offer assistance with placing students at the appropriate level, but in mathematics, research demonstrates a troubling predominant focus on practice items, in the design and/or classroom use of digital resources, on practice items (Murphy et al. 2014; Bolt et al., 2010). Some studies show diminishing effectiveness of treatments at higher grades (Ysseldyke & Tardrew, 2007). Nonetheless, some digital systems, if implemented with fidelity, appear to lead to improved student outcomes--but typically measured only by coarse-structured achievement tests or only through annual testing (Ysseldyke & Tardrew, 2007; Murphy et al., 2014; Bolt et al., 2010, Pane et al., 2014).

Two elements are sorely needed to provide teachers with research-based ways of monitoring and supporting students' ongoing learning progress: 1) a clearly articulated, developmentally appropriate, and appropriately fine-grained set of learning targets, and 2) informative and aligned measures of student progress toward the attainment of the learning targets. College and Career Standards accompanied by annual testing may provide a degree of guidance, but this combination is too general--and the assessments far too removed from day-to-day instruction--to effectively meet the two critical needs just identified. However, even assessments aligned to these standards still fail to communicate particular detail about learning. Templin and Bradshaw refer to the problem of the "subjective reality of unidimensionality" which results from the common test development practice of systematically favoring test items that fit a single, psychometric dimension of mathematical ability, rather than specific dimensions (2014), resulting in scales that are too general. Our recent innovations at the SUDDS research group to represent learning trajectories (progressions), have made us well equipped to provide the two elements noted above.

Competency-based learning and personalization are quickly becoming the models of choice for meeting diverse student needs across the country (Achieve, 2015; Tanenbaum et al., 2013). These two models complement one another: competency-based approaches aim to structure instructional decisions around student competency in obtaining clearly delineated objectives, and personalization models, among other aims, seek to create means by which students have personal agency in choice of how to reach those competencies. Competency-based learning (also called *proficiency-based*, *mastery-based*, or *performance-based* learning) has five features: 1) students advance upon demonstrated mastery, 2) competencies include explicit, measurable, transferable learning objects that empower students, 3) assessment is a meaningful and positive learning experience, 4) students receive timely differentiated support based on learning needs, and 5) learning outcomes emphasize competencies that include application (Patrick & Sturgis, 2015). Because competency-based learning requires students to progress based on their mastery of explicit, measurable learning objects including applications, a delineation of learning trajectories is required to identify those learning objectives (Daro et al., 2011; Hess & Kearns, 2011; Mosher, 2011). For this express purpose, Achieve, Inc. convened a meeting in May 2015 of representatives of 12 states partnering with Achieve to implement competency-based approaches, along with experts in the learning sciences. The resulting report (Achieve, 2015) argues for learning trajectories to help students get "deeper conceptual learning and application of knowledge" and that teacher are "equipped to connect day-to-day instructional planning with annual learning goals and longer-term goals that extend [across the years]." (p. 1)

Competency learning is often accompanied by or embedded within a second approach, especially in the context of digital technology, known as “personalized learning.” A personalized learning approach aims “to tailor learning to students’ strengths, needs, interests, and experiences” (Patrick & Sturgis, 2015, p. 11). It includes some type of customized learning plan, data on student knowledge and skills, interests, reflective knowledge of themselves as learners (Patrick & Sturgis, 2015), and has five characteristics: 1) learner profiles, 2) personalized pathways, 3) flexible learning environments, 4) an orientation to college and career readiness, 5) competency-based progressions (Pane et al., 2015; Patrick & Sturgis, 2015; AIR 2013). National Media Consortium’s Horizon K-12 Report identified personalization as driving promising new technologies (Johnson et al., 2014) and noted two paths of development, one driven by and for the learner to select personalized pathways and preference, and the other, by the machine intelligence, interpreting data and determining an adaptive pathway. Recognizing the critical role for “competency-based progressions,” this model also demonstrates a critical need to transform a widely distributed and uneven research base on learning trajectories into a set of specific tools and resources to integrate the knowledge base to support diverse students in achieving mathematical competence in personally meaningful ways.

Very little research has been conducted around these models, due to the recency of the approaches and their constant transformations. Research that has been conducted shows three promising trends: 1) significant increases in growth of knowledge, 2) particular successes with students (often underserved populations) at the lower performance levels, and 3) increased effects over time (Pane et al., 2015; Ready, 2014) The RAND study of 62 largely urban, charter, and district schools reports an effect size of .27 in mathematics across all students (K-12), in favor of personalization, with a smaller but significant effect size of .14 for the middle grades students in the study. In a two-year study of personalization at seven and then fifteen middle schools using the School of One approach (now Next Generation Schools), Ready (2014) reported mathematics gain scores of 15 % over national norms in year one and 47% in year two. Gains were more apparent for the students in the lower end of the performance scale. In spite of these promising early findings, a weakness with these particular studies is their reliance on a single outcome measure, the NWEA Measure of Academic Progress (MAP), a K-12 adaptive test that is based on a unidimensional growth scale (Northwest Evaluation Association, 2000). The strength of this measure is that the cross-grade scale allows for comparison across different timescales and grade levels, the ability to be given multiple times, and reporting on how long students spend taking the test. However, this test’s validity at upper grades, especially for students on grade level, is not well established, and it does not provide specific feedback on particular mathematical concepts learned. For example, although these studies report effect sizes, they are unable to report which particular mathematical concepts or skills that students in either group did or did not learn. We argue that this is where diagnostic learning trajectories-based measures could provide far more specific information on learning. It allows us to track progress across years, and it is designed to do so at those critical middle grade years where U.S. students often lose competitiveness internationally (Beaton, 1996). We further argue that the need is urgent for further development of alternative measures carefully tied to research on student learning: millions of dollars are being used in pursuit of these new models of competency and personalization, and we need a better means to know, in detail, if those investments are sound and, if so, how they can be improved.

A Next Generation Digital Diagnostic Assessment Measure based on Learning Trajectories

Our research team is well situated to contribute to, refine, and strengthen a new model of assessment, because of our understanding of both commercial and research initiatives. The approach we describe in this proposal is supported by our prior work funded by the Bill and Melinda Gates Foundation to create an open source digital learning system (DLS) for grades 6-8 mathematics (Confrey, 2015). The front end of the DLS is a learning map of nine big ideas in four “fields,” and is designed to help users examine learning goals at different levels of specificity and on different time scales. For example, the big idea “Position, compare, and operate on one dimensional quantities” is a high level learning goal that takes multiple years to master. In this region, though, users can zoom in to see related learning clusters (RLCs) that students must develop in order to gain proficiency in this big idea. One can zoom in further to see the learning trajectory (and related standards) associated with one construct to understand how incremental changes in students’ observable actions can indicate progress across days and weeks.

The back end of the DLS is comprised of diagnostic assessment items, an automatic reporting system, and analytics generated on various parameters of the DLS’s use (see Figure 1). Students access assessment items at the RLC level that have been designed to indicate their particular levels in the learning trajectories. Students navigate the map to explore links to open-source curricular materials, take assessments to gain insight into their progress in learning particular ideas, and engage with reporting

features--both to interpret their assessments and to collaboratively prepare for future learning with other students and with teachers. The approach leverages a key assumption: that the solution to effective personalization in competency-based learning is to successfully recruit students into more active roles in planning, managing, and assessing their own learning.



Figure 1: Components of a digital learning system (revised from Confrey, 2015)

We seek support to continue this line of research and the implementation of a next-generation assessment in 1) a school and district partnership with one high-performing district with leadership mentored in competency-based learning and personalization, and 2) with one new middle school, which serves a large percentage of students on free and reduced lunch, and whose dynamic leadership team is much earlier in the process of implementing digital learning and personalization. Starting from our current work in which we have articulated the underlying learning trajectories for middle grades into a set of proficiency levels, we are currently field-testing over 600 assessment items designed to measure students' conceptual understanding of the proficiency levels (with 1-2 items per level). By the beginning of the proposed project, we will have collected and analyzed, from students in 6-8th grade, more than 100 responses for each assessment item for each of the 404 levels within the 64 constructs across the 24 RLCs. From this base, the proposed work will systematically strengthen our knowledge of learning over time, including how to gather, interpret, and apply that knowledge to improve classroom practice and to further develop productive conceptions of competency-based personalized learning. To achieve these goals, we propose to address the following research questions:

Research Question 1: *How can we provide students and teachers with prompt, valid, and accurate diagnostic information on student progress related to nine big ideas using learning trajectories? What kinds of assessment prove most informative?*

Research Question 2: *How do students and teachers interpret and make use of diagnostic information in real time? What kinds of choices do these interpretations lead them to make about future learning?*

The **intellectual merit** of the work lies in pioneering a new form of assessment, of students' learning big ideas over time (Peck and Briggs, 2015), which provides precise and actionable information to students and teachers in real time. Further, because the diagnostic assessment system is within the context of an open source learning system based on research on learning trajectories, the project further investigates how this information can be used to transform and improve instruction. This research will contribute significantly to: a) investigating the degree to which promising approaches around competency and personalization result in solid cognitive outcomes for students with a different outcome measure than the MAP test, 2) advancing our knowledge about how different item types contribute to assessing LT progress, 3) advancing measurement modeling techniques to assess growth over time in ways that are personally meaningful to students, and 4) guiding teachers and students to interpret and act on the results of diagnostic assessments.

Results from prior NSF support

Confrey and Maloney were awarded a grant DRL-1118858, "Completing, Validating, and Linking Learning Trajectories for K-8 Rational Number Reasoning Tied to the Common Core Standards," starting

August 1, 2011 and expiring July 31, 2015, with a no-cost extension to July 31, 2016. The project has accomplished and continues work on a variety of goals:

Intellectual Merit (chronologically ordered from earliest to latest)

- developed the website TurnOnCCmath.net, which identified 18 LTs from K-8, provided an extensive set of descriptors of related CCSS-M standards, and synthesized the related research;
- completed proficiency matrices for all 18 LTs with related assessments (licensed to Amplify Education, Inc.) for major international middle grades curriculum;
- Confrey, while at Amplify as Chief Mathematics Officer, conducted related studies on sixth grade ratio, number, percent, and statistics;
- Confrey won Innovator of the Year award for NCSU, the first ever by an education professor.
- Confrey served on the Future of NAEP Panel, bringing progressions to the Committee's attention;
- Received and approved multiple requests from scholars to use the LTs in their studies
- secured support for a related project "Learning Maps Based on Possible Paths through Relational Learning Clusters (RLCs)" from the Bill and Melinda Gates Foundation to generate a new map based around relational learning clusters, big ideas, and learning trajectories, allowing the research team to build the DLS to assist schools in implementing LTs in relation to CCSS-M;
- developed LTs and related assessment items corresponding to 64 constructs, and is currently conducting field tests of these items;
- developed an active website, a blog for teachers (Teacher Connect), and access to the map.
- conducted design studies at NCSU on sixth-grade students' statistical reasoning; and
- recently received requests from two private companies, a large publisher's research division (Pearson) and smaller start-up, (Branching Minds) to use our various forms of LTs.

Broader Impact

- The TurnOnCCmath.net site has been used extensively both nationally and internationally. Presentations were given in nine states (NC, KY, CO, WA, MI, WV, ME, NJ, TX). The user base is tens of thousands. The Teacher Connect site has more than 200 teachers and 2000 students.
- Confrey and Maloney hosted a national meeting on diagnostic assessment with support of the Pearson Foundation and produced the report: Confrey, J., Hasse, E., Maloney, A., Nguyen, K. H., & Varela, S. (2011). *Designing Technology Enabled Diagnostic Assessments for K-12 Mathematics*.
- Confrey hosted a meeting for CCSS-M writers and other LT researchers to help provide writers a solid understanding of the current research base before CCSS was written.
- Confrey served on the CCSS National Validation Committee.
- The previous version of Learning Trajectories are currently licensed to Amplify Education Inc. for use in curriculum and assessment development. For two years, PI Confrey led the middle grades math development at Amplify as Chief Math Officer, and some of her doctoral students continued working on the product when she returned to NCSU to continue the research.
- Maloney and Confrey ran two MOOCs on the Equipartitioning LT.
- Confrey & Maloney presented a seminar on LTs in math for program officers at NSF (Feb. 2015).
- Confrey attended the May 2015 Achieve meeting, "The Learning Progressions in Competency-based Pathways" as an expert on learning trajectories with 12 states in Achieve's "Competency Based Pathways State Partnership."
- Confrey presented on the research at NCTM 2015 and NCSM 2015 and is scheduled to present it at AERA, 2016 and NCSM, 2016.

Publications

Books: Maloney, A. P., Confrey, J., & Nguyen, K. H. (Eds.) (2014). *Learning over time: Learning trajectories in mathematics education*. Charlotte, NC: Information Age Publishers.

Dissertations: Nguyen, K. (2010): Investigating the Role of Equipartitioning and Creating Internal Units in the Construction of a Learning Trajectory for Length and Area. Corley A. K. (2013): A Design Study of Co-Splitting as Situated in the Equipartitioning Learning Trajectory.

Refereed Journals: Prediger, S., Gravemeijer, K., & Confrey, J. (2015). Design research with a focus on learning processes – an overview. *ZDM—The International Journal on Mathematics Education* 47(6).

Confrey, J., Jones, R. S., & Gianopoulos, G. (2015). Challenges in Modeling and Measuring Learning Trajectories. *Measurement: Interdisciplinary Research and Perspectives*, 13(2), 100-105.

Confrey, J., & Maloney, A. P. (2015). A design research study of a curriculum and diagnostic assessment system for a learning trajectory on equipartitioning. *ZDM—The International Journal on Mathematics*



Education47(6).

Confrey, J., Maloney, A. P., & Corley, A. K. (2014) Learning trajectories: a framework for connecting standards with curriculum. *ZDM—The International Journal on Mathematics Education* 46(5): 719-733. Editorship: Prediger, S., Gravemeijer, K., & Confrey, J. (Eds.) (in press 2015). Special Issue on Design Research in Mathematics Education. *ZDM—The International Journal on Mathematics Education* 47(6). Book Chapters: Confrey, J. (in press 2015). Designing curriculum for digital middle grades mathematics: personalized learning ecologies. In Bates, M. & Usiskin, Z. (Eds.), *Mathematics Curriculum in a Digital World*. Charlotte, NC: Information Age Publishing.

Arnold, P., Confrey, J., Jones, R., Lee, H., & Pfannkuch, M. (in press 2015). Learning trajectories in statistics. In Pfannkuch, M., & Ben-Zvi, D. (Eds.) *The first international handbook of research in statistics education*.

Confrey, J., (2012). Better measurement of higher-cognitive processes through learning trajectories and diagnostic assessments in mathematics: The challenge in adolescence. In Reyna, V., Dougherty, M., Chapman, S. B., & Confrey, J. (Eds.) *The adolescent brain: Learning, reasoning, and decision making* (pp. 155-182). Washington, DC: American Psychological Association.

Conference Proceedings/Briefing Books: Confrey, J. (2015). The value of learning maps and evidence-centered design of assessment to educational data mining. Briefing book for Computing Research Association workshop, “Advancing Data-Intensive Research in Education.” Washington, DC, June 2.

Evidence of Research Products. The TurnOnCCMath.net is an open resource for researchers and teachers that summarizes research connected to 18 LTs from grades K-8. The SUDDS learning map is an open-source website with a Creative Commons license.

Relationship of Prior Work to Current Proposal

This proposal builds directly on two elements of our DLS: the learning map and the diagnostic assessments linked to LTs. We provide an example of each, to demonstrate how extensively the research on student learning influences the map and its learning trajectories, and how the overall construction carefully addresses the complexity and interrelationships of fundamental ideas in multiple ways.

Learning Map. The learning map is hierarchically structured around nine *big ideas* located in four *fields* (number; statistics and probability; geometry and measurement; algebra). For each big idea, 1-5 *relational learning clusters* (RLCs; Confrey, 2015) are identified (a total of 24 RLCs). An RLC is a set of 1-4 *constructs* that should be learned in relation to each other. Their configuration on the map shows the suggested sequencing within the clusters, from bottom to top; constructs positioned at the same “latitude” signal that they can be sequenced in either order. Finally, for each of the 64 constructs, a cross-grade *LT* is specified (displayed from bottom to top, with grade level associations to the right), which describes what is known from empirical study about significant student behaviors likely to surface as students progress from naive to more sophisticated understanding. These construct-mapped LTs are at the heart of the work.

Figure 2 shows the top level of the map displaying the four fields and the nine big ideas. Figure 3 then shows the big idea “Position, Compare, and Operate on One Dimensional Quantities,” zoomed to the RLC “Operating with Integers, Fractions, and Decimal. A user has clicked on the construct “Multiplication and Division of Positive Rational Numbers” (at lower right), to reveal its learning trajectory stack (right) and associated CCSS-M Standards (left). Other constructs in this RLC are “Addition and Subtraction of Positive Rational Numbers” and then “Operating with Positive and Negative Numbers” (which locates integer operations in 7th grade) and “Representing Large and Small Numbers with Scientific Notation.” This RLC includes elementary, sixth, seventh and eighth grade topics. Please visit www.sudds.co to view the live version of the map. Enter the username: demo@sudds.co and password: sudds and click “Login.” (These instructions will preserve users’ anonymity.)

LTs and Diagnostic Assessments. The proficiency levels for the construct’s LT (table 1) capture research on multiplication and division (Harel & Confrey, 1994) about how students learn the ideas while the associated standards only list the required concepts. Level 1 of the LT’s first level addresses GCF and LCM and the item is non-standard in that it asks for possible missing values allowing for multiple selections. Level 2 tests for a common misconception—the failure to insert a zero when the value brought down is smaller than the divisor (Ambrose, Baek, & Carpenter, 2003). Level 3 addresses multiplication in the context of arrays and Cartesian products, a topic that is reinforced in the probability constructs (Hiebert & Behr, 1988; Vergnaud, 1994), and illustrates multiplication as “referent composing” (Confrey, 2012). From arrays, students are moved to area models for multiplication of fractions which is extended to mixed numbers (Izsák, 2008). The items check students’ application of the distributive property (Steffe, 1994). The inclusion of arrays and area models may seem counter-intuitive within a big idea devoted to one dimensional representations, but research has shown these models to be cognitively more intuitive for beginning

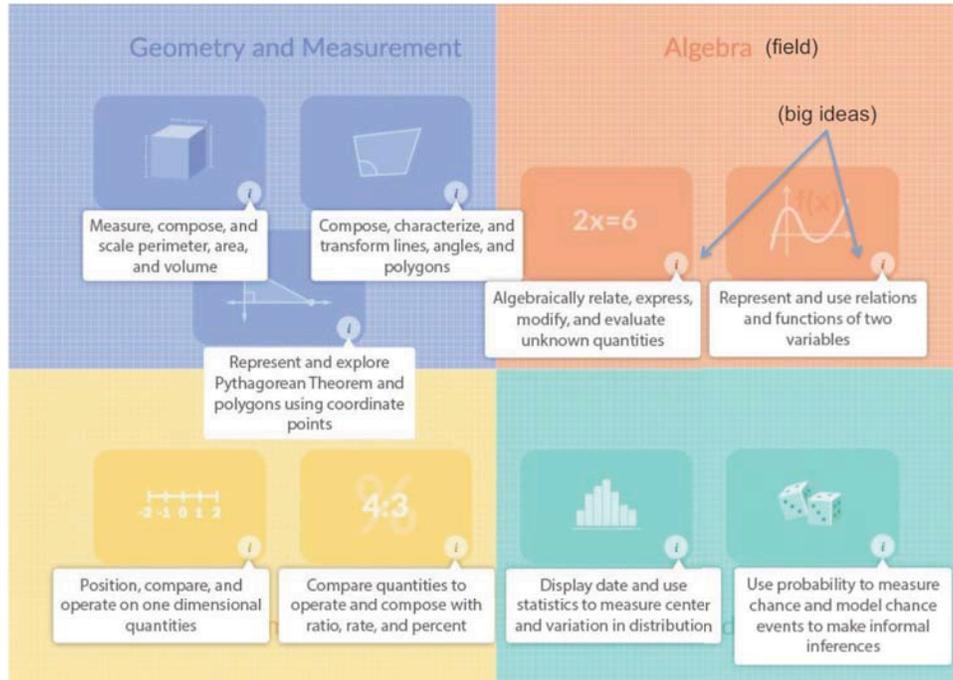


Figure 2. Top level of Learning Map, grade 6-8 mathematics: the nine big ideas



Figure 3. Standards and Learning Trajectory associated with one construct.

students to learn multiplication (Russell, 2012; Hiebert and Behr, 1988) and the learning trajectory is designed to move students from these models to the more abstract and generalized model of scaling which is “referent preserving” (Confrey, 2012; Otto et al., 2011) in level 6. Level 7 explores the misconceptions that multiplication makes bigger, division makes smaller (Greer, 1992; Greer, 1987; Bell et al., 1984). Level 8 examines division of fractions in the context of measurement (Dougherty, 2008; Simon & Blume, 1994) and level 9 checks for understanding of the inverse relationship of multiplication and division for rational numbers. Level 10 investigates students’ reasoning about moving the decimal point in multiplication.

Table 1: Proficiency levels (from bottom to top) for a learning trajectory and sample assessment items

Proficiency Level for "Multiplication and Division of Rational Numbers"	Sample Diagnostic Assessment Item
<p>Level 10</p> <p>Divides decimals by shifting decimal point and dividing, and explains</p>	<p>Matthew multiplied the number A times the number B and got the answer 1196. Then he moved A's decimal point one place to the left and B's decimal point two places to the left and multiplied the resulting two numbers to get a new answer. Which of the following best describes Matthew's new answer? (Students are provided a set of choices.)</p>
<p>Level 9</p> <p>Divides fractions by multiplying by the inverse (reciprocal) and explains</p>	<p>An inventor created a shrinking and magnifying machine that makes things smaller or larger than their original size. For example, he shrinks something to half its original size by setting the machine to $\frac{1}{2}$. He triples an object in size by setting the machine to 3. The inventor changed the size of four objects to test his invention. (objects are given)The inventor wants to each object back to its original size. What setting should the inventor use for each object?</p>
<p>Level 8</p> <p>Models division of fractions (a/b)/ (c/d) using measurement model (how many c/d in a/b ?)</p>	<p>For a running race, Stephanie is marking a $5\frac{1}{4}$ km route with rope. The rope comes in spools, and each spool contains $\frac{7}{20}$ km of rope. How many spools will it take to mark the route?</p>
<p>Level 7</p> <p>Multiplies decimals and justifies in relation to the context and/or model</p>	<p>Mr. Xi wrote the following examples on the board: 0.3×0.07 1.3×0.8 0.2×4.1 7.5×2.1 Which example has an answer smaller than both of the decimals? Which example has an answer larger than both of the decimals?</p>
<p>Level 6</p> <p>Converts sequences of multiplication and division of whole numbers into multiplication by a single rational number and computes</p>	<p>Alicia has a copier that can change the size of a photograph. She can multiply by a whole number to make the photo larger and divide by a whole number to make it smaller. She can also use a combination of multiplying and dividing by whole numbers to resize the photo. The student is asked to predict the relative size of the photo after sequences of scaling.</p>
<p>Level 5</p> <p>Applies and gradually extends area model from multiplication of proper fractions to improper fractions and mixed numbers, identifying 1 square unit and using distributivity</p>	<p>Reynaldo was asked to solve the following problem: $2\% \times 4$ Ya. He solved it by multiplying $2 \times 4 = 8$ and $\% \times Ya = \frac{3}{40}$. He gave his answer as $8\frac{3}{40}$. The answer options ask the student to agree with a variety of possible interpretations.</p>
<p>Level 4</p> <p>Applies and gradually extends area model from two whole numbers to multiplication of proper fractions, identifying 1 square unit as a composed unit</p>	<p>Marcie is buying fabric for her Halloween costume. She is looking at a shelf with leftover pieces for sale that are all different sizes. The price is based on the area of the piece and is \$10.00 per square yard. Below are the dimensions for several different pieces, measured in yards. If Marcie has \$5.00, which of the pieces can she buy? Check all that apply. Choices: A. $\frac{1}{2} \times \frac{1}{2}$ %, B. $\frac{1}{2} \times \frac{3}{4}$ %, C. $\frac{3}{4} \times \frac{1}{2}$ %, D. $Ya \times \frac{1}{2}$ %, E. $\frac{1}{2} \times \frac{1}{2}$ %</p>
<p>Level 3</p> <p>Applies and extends array model for multiplication to Cartesian products (three shirts and two pairs of pants to create six outfits), recognizing composed units (outfits)</p>	<p>A dog walker walks three different dogs each day and only walks one dog at a time. There are four routes that she uses to walk the dogs. How many different combinations of dogs and routes can she take before she travels the same route with the same dog again?</p>
<p>Level 2</p> <p>Divides multi-digit numbers using the standard algorithm</p>	<p>This is Bernie's work to divide 211 by 20. Review Bernie's work for correctness, then enter the correct answer to this division problem below.</p> $\begin{array}{r} 10.55 \\ 20 \overline{) 211.0} \\ \underline{20} \\ 110 \\ \underline{100} \\ 100 \\ \underline{100} \\ 0 \end{array}$
<p>Level 1</p> <p>Factors, prime factors, and finds GCF and LCM</p>	<p>The greatest common factor of {24, 36, 72} and the least common multiple of {4, ?} are the same number. Which of the following numbers could be the missing number in the second set? Check all that apply. Choices: 1, 2, 3, 6, 12, 24</p>

After taking a diagnostic assessment, a student (and the teacher) receives a report that provides feedback on the percentage of items correct for each construct, a scaled score for the overall RLC, and a table showing performance on each item (see figure 4). Students can review their submissions and submit explanations directly to the teacher for possible rescoring. (The only test currently embedded in the map is in the first statistics RLC, since other assessments are currently aligned directly to specific curricula of partner schools. To view the sample test, log in to the map (see above), click the histogram panel in the Statistics and Probability field, then click the check-mark in the "Displaying Univariate Data" RLC)



Figure 4: Example of a diagnostic assessment report at item-, construct-, and RLC-level for introductory statistics (gr. 6).

During Spring 2016 (our final period of funding for our current NSF research), our school partners will complete the field-testing of the diagnostic items (item writing will be completed by 12/31/2015). During the spring, the team will analyze the data to calibrate and revise items as needed. The items will be calibrated using item response theory, and unidimensional scales for each LT will be generated in preparation for psychometric modeling activities. Discussions with practicing teachers and students about the diagnostics and reporting will also be undertaken for our final report. During summer, our partners will align their curricula to the learning map, and the diagnostic assessments will be deployed to the map.

Research and Development Design

Theoretical Framework

The major theoretical framework that underlies the proposed new work is that of learning trajectories (LTs). LTs are articulated descriptions of empirical behaviors that describe landmarks and obstacles likely to be encountered by students as they progress from naïve to more sophisticated conceptions (Simon, 1995; Confrey, et al., 2009; Clements and Sarama, 2014; Lehrer & Schauble, 2015). Rather than to conduct a logico-mathematical analysis of the structure of an idea, a learning trajectory describes an elaboration about how ideas develop in the minds of students, when they are given compelling tasks involving active student participation in the learning process. LTs incorporate misconceptions, systematic errors, alternative conceptions, and strategies of increasing generalization.

LTs are leveraged in this project in two ways. First, they are tied to the big ideas through the RLCs, which group together constructs that need to be learned interactively. Because each construct has an underlying trajectory, teachers use these to organize their instruction and to recognize, value, and build on student thinking. Secondly, they are tied to assessments. Assessments have been built for the vast majority of levels in the LTs using an evidence-centered design (ECD) framework (Mislevy, 2006). A first round of assessments, later revised and strengthened by researchers, were written by classroom teachers; after a

week of writing to proficiencies, teachers expressed they: a) finally understood the standards, b) could see how to target their instruction to serve students better, and c) knew better how to support discourse.

The approach to the research is through an “assessment for learning” process (Black, 2004) rather than directly through the examination of student thinking. Briggs and Peck (2015) challenged the assessment community to undertake the study of conceptual development over time, using Confrey’s work as an example. In their article, they argued for “a new approach to the design of vertical scales that is premised upon a priori hypotheses about growth in the form of a learning progression hypothesis [as a means to **create**]. a meaningful criterion-referenced interpretation of growth magnitudes – from a coherent conceptualization of what it is that is growing over time” (p. 80). In a response to Briggs and Peck’s article, Confrey, Jones, and Gianopulos (2015) agreed and emphasized the need for such assessments to occur right after the relevant instruction.

Study Goals

The goals of this project are to: 1) fully develop a middle grades diagnostic assessment system built around learning trajectories, 2) provide evidence of feasibility, utility, and promise from new rounds of field-testing in collaboration with our district and school partners, and 3) document ways to use the information to meet the diverse needs of students. By providing a means to measure progress along progressions, towards big ideas, over time, we will offer the field a way to examine what is actually being learned within the variety of approaches to personalization and competency-based learning. Such evidence will inform not only further studies to examine the effectiveness of such approaches, but will open up the possibility of assisting teachers in ways to differentiate instruction to meet the needs of all students while still using methods that support small group work and discourse (Freeman et al., 2014).

In order to create a full diagnostic assessment system for learning around learning trajectories we will develop, research, and refine a full validation argument for the use of our tools. In the context of diagnostic assessments, such an argument will ensure that we address concerns for validity and reliability within the context of use. This argument will span the full range from addressing how teachers and schools target the learning of the nine big ideas over a time span of three years, assist students to develop a mindset for consistent progress and self reflection, and show documented learning along the learning trajectories. We propose the following four design and implementation components to reach our goals:

Component 1: Adding and contrasting and comparing varied forms of diagnostic assessments

We will use the data from field-tests to guide us in how to build out the diagnostic item database to support fuller use. A priority will be to replace items that do not seem to be performing well in terms of measuring the trajectory or exhibiting other qualities that bias student response. This includes a review to ensure items are fair, culturally and linguistically sensitive, and grade appropriate. In addition, we will quickly and cheaply create clones for existing items when appropriate by varying the number magnitude, objects, or verbs in the problem, but keeping the clone at the same level as the original item. To facilitate this work, we will develop a protocol for field-testing new items within the cycle of diagnostic assessment use. Rather than sequester “secure” items to distinguish real and practice tests, we prefer to enlarge the existing item pool to support students taking multiple tests from the same pool. The next priority is to broaden the types of items we can support to include natural extensions of current item types. This will involve, for instance, adding drop-down menus to the current set of data entry buttons, giving item designers control over the the number and location of data entry buttons, including mixing data entry boxes/buttons within the same item, and giving users the ability to answer questions using an equation editor. Because our map and assessment system will be open source, others will be able to use these resources to build their own assessments.

Another type of item is called “item generation environments” (IGEs; Confrey & Maloney, 2012). These are configurable representations and tools designed to dynamically capture students’ problem-solving processes and strategies. Our first IGE, a dynamic number line, is in final development stages. It can be provided as an open tool for students to use in problem solving, initially presenting only a bare number line and zero. Students build the representation using the tools of equipartitioning, replicating, and scaling. The number line on the bottom can be configured to show fractions, decimals or irrationals and put them in lowest terms, mixed numbers, etc. Artifacts can be placed on the number line to identify points as needed. The number line can also be preconfigured to certain conditions by the item writer, and placed into an assessment; it then collects strategy and process data on student work. Determining how to return these data to users is a next step in that process, to be completed this spring. In a recent round of “playtesting” (an agile method of trying out prototypes with students in our lab), students were asked questions such as: Using as few tick marks as possible, create tick marks that have two, three, and four decimal places the between 0.9 and 1.0. Our initial results suggest that students find the task of creating a number line more

challenging than locating a point on the number line. These results are consistent with research from the field (Saxe et al., 2007). During development of diagnostic items and elaboration documents for each LT, we have identified top candidates for building IGEs. If funded, we will review these, select the best for prototyping and playtesting, and develop them for use and research with the overall diagnostic system.

In this first component of the research, our goal is to systematically investigate the questions: “What are the changes in what is learned about student thinking when using different item types and when examining process and strategy data?”, “If you vary the measures of specific trajectories using different item types, how will the student outcomes and diagnostic information vary across different student populations, and what does this say about cultural and linguistic sensitivity?”, and “Do students and teachers find the data of more or less helpful and convincing depending on the item types?”.

Component 2: Investigating models for learning across time based on learning trajectories.

The second component is the development of measurement models that can provide valid and meaningful inferences about student competencies described in the LTs. This addresses two critical challenges facing the field of measurement: 1) how to relate assessment measures to meaningful descriptions of learning over time (Briggs & Peck, 2015; Confrey et al., 2015), and 2) how to select the right-sized unit of analysis for measurement (B. Gong, personal communication, August 2015). Item calibrations and conventional psychometric techniques from our analyses of these responses provide a starting point for modeling learning, but do not sufficiently address questions that require large numbers of students, longer time frames, or involve the relationships between diagnostics and subsequent instruction. Further, we will pursue advances in both psychometric modeling (see letter from Brian Gong, National Center for Improvement of Educational Assessment NCIEA, Advisory Board member), and in supporting students and teachers to use assessment tools to better understand student thinking.

Our envisioned assessment innovations pose two new measurement challenges. The first measurement challenge is to accurately model the hierarchical and multidimensional structure of the 64 LTs nested in 24 RLCs across three grades. This is an issue that researchers have only recently begun to study, and multidimensional vertical scaling (across grades) has not been implemented widely. We have reviewed measurement literature to identify candidate statistical models that can meet the goals of the diagnostic testing system. A number of statistical models have been identified as strong candidates, including cognitive diagnostic models (CDM) (Tatsuoka, 2009; Rupp, A., Templin, J., & Henson, R., 2010; Henson, Templin, Wilse, 2014), unidimensional item response theory models (IRT), structural equation modeling (SEM), multidimensional IRT models (MIRT), including its application to a learning progression for statistical reasoning (Lehrer, Kim, Ayers, & Wilson, 2014), the bifactor model (Reise, 2012), the two-tier MIRT model (Cai, 2010), and the hierarchical MIRT model for cognitive diagnosis (Hansen, 2013).

Each type of model serves a different purpose. IRT methodology will permit us to create provisional unidimensional vertical scales for each learning trajectory. Within each LT scale, CDMs will be used to model mastery profiles that attempt to explain why a student has or has not mastered a given item or task. Simple CDMs and unidimensional IRT models have small to modest sample size requirements, making them the most likely candidates for the first stage of calibration and scaling. As the user-base of the DLS grows and sample sizes increase, we will be able to turn to more complex and advanced models, such as SEM and MIRT to model the structure of RLCs. The choice of model is influenced not only by sample size, but also by the cognitive structure of the RLC, the number of dimensions or attributes, and how the attributes relate to one another, the structure of the test blueprint, and the desired level of transparency. As sample sizes permit, we will evaluate the internal structure of each cluster assessment, using a combination of exploratory and confirmatory factor analysis to identify and remove nuisance variables and determine the most appropriate model to represent each RLC. To detect and remove items with possible bias, differential item functioning analyses will be conducted, based on student ethnicity, gender, and socio-economic status. Rival models will be compared using model fit statistics, the ratio of variance accounted for by the cluster factor over the constructs (LTs), the extent to which cluster scores indicate a common variable, and the reliability of construct scores. In keeping with our goal of developing many distinct scales that are each sensitive to each learning trajectory, we will also explore the use of Bayesian Nets (Almond, Mislevy, Steinberg, Yan, & Williamson, 2015) to model the relations among the LT scales.

Assessments with multiple item types, item clones, and IGEs pose a second set of modeling challenges. Each IGE has multiple ‘design variable settings’ that control and change the nature of the items, including the difficulty of the item and the alignment of the IGE to the stack levels. The many design variable settings lead to a great number of combinatorial possibilities. We believe we can address this concern by experimenting with IGEs in the practice tests, and can identify the most promising combinations of design

variable settings that can be used to generate a manageable number of IGEs which can be embedded in the operational tests for field-testing. The practice test settings will also provide opportunities to investigate different scoring algorithms for use in the operational assessments. If managed carefully, the number of these constrained IGEs would be small enough to yield a sample size sufficient for calibrating with an IRT or MIRT model. Once these constrained IGEs have been calibrated they will be treated as other tasks and items within the assessment. Once sufficient data are collected, another possibility is to use an appropriate regression model to predict the IRT or MIRT parameter estimates using the design variables. Assuming we obtain sufficiently precise and accurate predictions, we can substitute actual item parameter estimates with predicted item parameter estimates. This might permit us to use IGEs with fewer constraints while maintaining calibrated items needed for scaling purposes.

Component 3: Studying teacher and student interpretations of data

This component comprises related issues of teacher and student understanding of assessment data in relation to the underlying LT framework. The first issue is a challenge in competency-based and personalized learning: for teachers to understand the LTs well enough to interpret the students' diagnostic assessment data. Teachers often insufficiently evaluate students' thinking due to the paucity of discipline-specific frameworks for making sense of student thinking (Borko et al., 1997; Chazen & Ball, 2001). Learning trajectories hold much promise to provide the needed interpretive framework (Hess, 2008), e.g., learning trajectory-based instruction (LTBI; Sztajn et al., 2007), but teachers often focus on the endpoints of trajectories in the assessment results, and ignore consequential intermediate states of student thinking (Kim & Lehrer, in press). William James's (1971) adage, "A difference which makes no difference is no difference at all," applies here: levels of LTs may seem too fine-grained to teachers until they see how the data (patterns of performance on assessments) correspond to levels of the learning trajectories. Our conjecture here is that the assessment data will lead teachers to a deeper understanding of the LTs. By grappling themselves with students' struggles in the LT-based assessments, teachers will more deeply understand the distinctions the LT levels make about student learning growth. This was confirmed as a highly productive strategy for teachers to learn about student thinking, in work conducted last summer. To explore the validity of the conjecture, we will periodically ask teachers to explain the distinctions among LT levels to see how these change over time.

To better understand what teachers need to know to understand assessment data, we will examine how teachers interpret the student data in relation to the LTs, using the framework of "data literacy for teaching" (Gummer & Mandinach, in press; Mandinach & Gummer, 2013). Data literacy for teaching is developed in an inquiry cycle comprising stages of framing questions, using data, turning data into information, transforming information into decisions. We will choose particular data sets from diagnostic assessments, and ask teachers, working in groups, to report on what they observe in, and how they interpret, the data, especially in relation to student learning over time. We will also provide item analyses, including information on students' choices of distractors. These investigations will involve particular patterns of data results, and different levels of aggregation

Also critical in this third component is the issue of representing assessment data to communicate differences in student thinking and support teacher reasoning about the data. For example— a) Do higher numbers on an assessment report scale suggest sufficient interpretations of the ideas students are developing? b) How should we represent the relative sophistication in different strategies even if final answers are the same? Investigations into the interplay between representations of assessment data and the interpretation of student thinking/growth by teachers (or entire departments) will inform revisions to the reporting mechanisms and the development of supports for a variety of types of analyses.

This component also bears on better understanding how students themselves can make use of LT based assessment data, and how the reporting system can best be designed to support students to interpret their own progress within a competency-based framework. LT-based assessments provide new interpretive challenges when students are accustomed to "percent correct." A touchstone for this work is the framing of "growth mindset" (Dweck, 2006), which asserts that intelligence is developed and struggle and mistakes play a central role in learning success. We will investigate how students can be supported to interpret the diagnostics results to evaluate their own learning and demonstrate their understanding.

During summer months of the project, we will conduct workshops, on interpreting assessment data with a learning trajectory perspective, and continue to study the ways in which teachers use these frameworks to inform their understanding of student thinking. In later years of the project we will study our refined reporting features in classrooms, observing students and teachers as they interpret assessment

data, focusing on the roles teachers and students play in collaboratively interpreting the data, and their claims about learning.

Component 4: Making instructional decisions based on data interpretation

Delineating how teachers and students act on the data is a key element of a full validation argument addressing the issue of consequential validity. We seek to know if diagnostic results tied to LTs can increase teachers' use of tasks that require meaningful mathematical inquiry, and promote increased scaffolding of students toward more sophisticated thinking (Stein et al., 2008). From our field-testing, teachers have reported using the assessment results in formative practices (Heritage & Popham, 2013; Wilms, 2003). The teachers assign groups of students to review items, then to explain to others why they gave certain responses and why those answers were correct or incorrect. Other teachers report they re-teach a topic and use the learning trajectories more explicitly by addressing the related misconceptions. Given this variety of responses, the first research approach we plan to undertake is to collect teachers' invented practices, using a form of teaching logs at our partner sites. Organized into professional learning communities (PLCs), teachers will develop and share, and we will document, best practices.

The second approach involves conducting department-level design studies (Stephan, 2015) "engineering particular forms of learning, and systematically studying those forms of learning within the context defined by the means of supporting them" (Cobb et al., 2003, p. 9). We will run at least one design study in partnership with our schools each semester. Teachers will plan curricular activities together, placing materials as links on the map. Students will take pre- and post-tests. Researchers will observe instruction both before and after students take the diagnostics, interview teachers to document instructional plan changes, and hold focus groups with students. Among the questions investigated will be: "How do the data interpretations from diagnostics affect learning of specific topics and ways of thinking?", "Do they make classroom instruction more "student-centered" and/or focused on "deep learning"?", "Do teachers use student responses more extensively, do they reference back to the learning trajectory levels, do they assign students to discuss and retake tests, and is there evidence of more focus on the conceptual underpinnings of ideas?"

The third approach will be to focus on how the data are used to differentiate via student grouping in order to meet the diverse needs of students. We will examine the use of diagnostics' results can be used by teachers and students to develop flexible grouping patterns, and answer questions such as: "How do teachers and students form and re-form groups based on the data?", "What combinations of collective traits (varied levels of performance) result in constructive learning?", "Are all students provided opportunities to work collaboratively?", "How feasible do optimization algorithms seem to generate recommendations for grouping students according to a given strategy, and can those recommendations inform (not replace) a teacher's decisions to group students?"

In summary, the combination of these four components will assist us in moving from a prototype DLS system to one ready for wider implementation and more rigorous testing of effectiveness. Because the four components span the assembly of the test to the use of the results, the study promises to provide insight not only into *what* works, but *why* and *how* it works.

Design and Implementation Processes

The project will be conducted at three sites selected for contrasting conditions: lab-based studies at NCSU, the Freehold School District's (NJ) two middle schools, and Highland Middle School in North Carolina. The lab studies bring students together for 1-3 week periods studying prototypes. iPads and an Apple TV support trials of new approaches. Our studies involve the three phases of planning, conducting and retrospective analysis (Cobb et al., 2003). All sessions are video-recorded, with artifacts of instruction collected for further study and use in professional development. We use findings to inform revisions to our designs and our research protocols.

Freehold Township School District in New Jersey was selected based on to its pioneering work on personalization of approach and competency-based initiatives. It has two middle schools with a total of 1500 students, 24 regular math teachers, and 12-16 special education teachers working in mathematics. Superintendent Dr. Ross Kasun, Assistant Superintendent Dr. Pamela Haimer, and Supervisor of Mathematics Dr. Charlene Marchese contacted us because they were undertaking a major reform in competency-based education, using 1-to-1 computing with Chromebooks to address personalization. Dr. Kasun is nationally recognized for his leadership in developing a digital infrastructure to support student learning in his district. He was one of 100 district leaders selected last year to participate in the "ConnectED to the Future" Summit at the White House, and was one of ten Superintendents selected to

offer input on the Future Ready leadership rubric. He and his colleagues had been using Confrey, Maloney, and Nguyen's prior work on learning trajectories (www.turnonccmath.net). Freehold is currently field testing assessment items and exploring cross-grade, project-based approaches to personalization. The school demographics data report 78% White, 4% African American, 8% Hispanic, and 9% Asian across the two middle schools.

Highland Middle School is a brand new school with 900 students and 11 mathematics teachers, located about one hour from Raleigh. Each student has access to a Chromebook and wireless internet. Highland Middle School is located in the Harnett County School District, a high-needs school district with 57% of students receiving free and reduced price lunch. The school's feeder elementary schools are very different in regards to students' proficiency: about half the students are from one of the lowest performing and highest poverty schools, and the other half from a top five high performing school with a large military population. The school demographic data report 50% White, 25% African American, 16% Hispanic, 7% Multi-racial, and 2% Asian and Native American. These schools have been field-testing our items and provide opportunities to work with students with diverse backgrounds and schools that address multiple educational challenges and goals.

This proposed study, in partnership with our districts, would be conducted over the course of three years. In the following sections we describe the research and development activities for each year.

Year 1: Fall 2016 - Summer 2017

During year 1 we conduct a thorough review of the item pool currently under development, to determine where item clones could be efficiently produced and develop guidelines for generating them. We add item-authoring capacity, and author three new item types: drag and drop, drop down menus, and items where data entry can be more flexibly located. We build three IGEs to measure student thinking in the fields of number and geometry. The item reviews for clones include ways to code items and define Q-matrices for each cluster to model the various traits of the items. We revise scoring rules as we define mastery attributes, fit response data from current field testing to CDM and MIRT models, and revise items as needed. During summer 2017, we conduct cognitive interviews with new item types.

Concurrently, we conduct a lab-based design study to investigate how students interpret score reports and work with teachers to pilot test a survey that will be used in years two and three to generate evidence of the kinds of decisions teachers make after interpreting assessment data in their classes from our system. These analyses inform our ongoing design and revisions of the student and teacher reporting mechanisms. During the summer of 2016 we work with select teachers from partner schools to align their curriculum with our learning map. During the summer of 2017, we conduct a professional development workshop with teachers on using the LTs and assessment data to inform instruction, and interview a subset of these teachers to identify the ways teachers conceptualize LT-based assessment.

Year 2: Fall 2017 - Summer 2018

By the second year of the project, our school partners are fully utilizing the assessment system and reporting mechanisms. As classrooms use our system, we embed new items from the year 1 clones, new item types, and IGEs for field testing purposes. We model and analyze these items, revise them, and gradually add them to the system as we better understand the item properties. We also design and build 3 IGEs for algebra and data and statistics. We more fully develop our measurement models as our sample of item responses grows and we incorporate what we learned from our year 1 modeling efforts.

In our partner schools, we seek to better understand how teachers and students are interpreting the data from the assessment system. We observe a sample of students as they read and discuss a report from a cluster test, and conduct interviews. We video two sessions of grade-level teachers collaboratively discussing data from assessments for all 3 grades and analyze these to better understand how teachers individually and collectively interpret data. We also select two teachers, one from each partnership, to observe across two clusters to learn about the decisions teachers and students make about instruction and how our system influences those choices. Finally, we administer a survey to all participating teachers at 4 points throughout the year to determine the kinds of choices teachers are making across all of our partner schools. During summer 2018, we conduct another teacher workshop focused on what teachers learned about assessment in the previous year and how to continue to grow in their assessment and instructional practices. We again interview teachers to gauge how thinking about assessment has changed in our partner schools.

Year 3: Fall 2018 - Summer 2019

The assessment system will be complete with all item types, clones, IGEs and reporting mechanisms for school-wide use. Teachers in partner schools begin their second year of instruction with

curriculum aligned to our map structure. We use the data from use in schools to continue to model, analyze, and understand the properties of the measurement tools in each RLC in the map. We revise items where needed, update item characteristic matrices, and revise competency definitions within the LTs and items. We also develop thorough documentation for the assessment system for users.

We draw on our findings from year 2 to revise our observation and interview protocols, and again observe students and teams of teachers as they interpret data. We administer surveys to all teachers 4 times throughout the year, and interview samples of teachers and students. Towards the end of the school year we conduct short interviews with all teachers to determine how they are thinking about assessment and LT frameworks for interpreting student thinking. These interviews also focus on the project as a whole and their experience with our assessment system.

Data Collection and Quality

All design studies will be video recorded by two cameras, one in the front of the room and one in the back. We use a variety of high quality microphones to capture audio data from teachers and students. During instruction we use cameras and iPads to collect images of classroom artifacts, and students participate as they record work products from their inquiry. These are all stored in secure NCSU servers.

Assessment related data are captured within the DLS and stored on a secure server. Per our IRB protocol, all student data remains anonymous to researchers. Unique student identification numbers (IDs) are created for each student and these IDs are used for analysis and reporting purposes. If any students have opted-out of the research work, their records are removed from the data files prior to data analysis. Our quality control procedures begin by merging student responses with test assignment forms to confirm that nearly all assigned tests were actually administered. Test data are accumulated across time and across forms. Descriptive statistics are generated for each item and any unexpected response frequencies or other anomalies are inspected. Test score distributions are visualized and any outlier scores are examined to confirm data accuracy. The data are also screened for response patterns that may indicate an incorrect key.

Analysis and Methods for Producing Findings

Our analysis techniques draw from two traditions, design research and psychometrics. For design studies, we debrief at the end of each day and review data to generate initial conjectures and revisions to the design. At the end of each study we conduct qualitative analyses of the video and images of artifacts both to inform particular questions, such as student interpretations of assessment data, and to identify new questions or areas for further inquiry. These data are coded and stored on our secure servers.

We analyze assessment data using a variety of psychometric techniques. These techniques do not only include statistical models, but also include cognitive interviews with small numbers of students around items and assessment forms, assessment use in design studies, teacher reviews and focus groups, and analyses of readability and structure which all lead to iterative refinements to the items. We model the item responses using a variety of approaches including classical test theory, item response theory, and latent class modeling. We use these to produce measurement properties for items, forms, and scales. From these combined data sources, our interdisciplinary team conducts a mixed method analysis. We regularly meet with partners to discuss progress, results, conjectures, and refine approaches.

Mechanisms to Assess Success of the Project

We will use an external advisory board to assess success in meeting project objectives. The advisory board chair, Dr. Dan Heck, will take responsibility for ensuring that the board's activities and recommendations are sufficiently rigorous and independent. The board will meet for 1.5 days during each year of the project to review, discuss, and give feedback to our ongoing research and development activities. with project summaries sent in advance The chair of the board will spend two days before each meeting with the research team to determine the type of information the advisory board needs to hear in order to provide sufficient accountability and guidance. After each meeting, the chair will review notes from the advisory board members, and draft a guide for the project's next year of work. The chair will also review the prior years work to ensure that the board's advice is being given appropriate consideration and action.

Advisory Board Chair

The advisory board will be chaired by Dan Heck, PhD. Dan is a vice president and senior researcher at Horizon Research, Inc. Dan's research has a long history of examining changes in classroom practices, and he has studied different ways teachers use data-driven practices to improve instruction.

Advisory Board Members

The Advisory Board consists of five individuals with diverse expertise in research, assessment, and learning in a variety of contexts. Dr. Brian Gong is the Executive Director for the Center for Assessment at the National Center for the Improvement of Educational Assessment. His Center has conducted extensive work on learning progressions, content standards, comprehensive assessment systems, growth and value-added measurement, report design, and the use of assessment information to inform instruction, evaluation, and accountability. Dr. Charlene Marchese is the Supervisor of Mathematics at Freehold Public Schools, one of our partner districts. She has worked in recent years to support teachers in her district to move towards more competency based and personalized instruction, and will provide insight into “on the ground” realities of our efforts. Scott McQuiggan leads the Curriculum Pathways division at SAS Institute, Inc. He leads an interdisciplinary team focused on the development of no-cost educational software and has expertise on analytics. Dr. James Pellegrino is the Liberal Arts and Sciences Distinguished Professor and Distinguished Professor of Education at the University of Illinois at Chicago. His research has long focused on providing teachers information that supports instructional decisions, and has more recently studies learning progressions in math curricula. Dr. Michelle Riconscente is the CEO at Designs for Learning, Inc. She was formerly the Director of Learning and Assessment at GlassLab and an assessment specialist at ETS. She headed the GlassLab work on learning maps and integrating data into useful reports.

Dissemination

We use agile methodologies and forms of social media to recruit early adopters of our assessment tools. We blog on our progress regularly and make curated, high quality, and free materials accessible to users. We are actively seeking an array of partnerships: both early adopters and dedicated users. We advertise online concerning our presentations at conferences. We also have a designated online community called Teacher Connect where teachers interact with our products as they are under development, communicate with each other, and provide feedback. We already have over 200 teachers members on Teacher Connect. We seek avenues to get the work to rural and urban schools through partnerships with organizations serving groups that are underrepresented in STEM. These often produce interesting collaborations and get the word out about our work in educational technology circles. We will continue to publish and participate in national and regional forums, using conventional academic methods as well, presenting invited talks and papers at NCTM, NCSM, PME, and CCSSO’s assessment meetings. In addition to these efforts, we will continue to collaborate with other research efforts that aim to develop assessment tools grounded in learning trajectory research for instance Gates/Hewlett’s Assessment for Learning collaborative. We also have partnerships with private enterprises such as Achieve, Pearson, ACT and others. In doing so we hope to build a robust community of researchers, practitioners, and innovators that has the potential to improve the state of assessment for public education.

Expertise of Research Team

The Principal Investigator spent two years in a corporate setting leading the development of a large middle grades curriculum project and learned to apply agile methodology. She developed methods to use lab-based studies to create prototypes while carrying out longer-term field tests. She returned to the university to pioneer ways to build agile and responsive designs in academic settings. The entire DLS and assessment system have been engineered *in under a year* by a talented, motivated, and hardworking team of nine people. The team includes an experienced psychometrician. Our UX designer has a PhD in neurology, applying her expertise to design. Our software engineers have experience in building games, designing for instruction and developing and supporting large databases. Our outreach coordinator spent two years building curriculum in a digital company.

Broader Impacts of the Proposed work

The potential for broad impact is substantial as the need for this measure is broadly recognized. As an open source enterprise with licensing to the Creative Commons, many different schools, regardless of wealth, can implement the approach. Possibilities to pursue commercial uses through licensing also exist, and the team is exploring a start-up company. Conversations about further partnerships are in progress through the Gates Foundation, the Teaching Channel, and other major players. The DLS is LTI compliant to link to learning management systems. Using responsive design, it is device neutral. Our work will be presented to district leadership in middle grades from across the state of North Carolina. Schools have expressed desires to access to the system as soon as available. Furthermore, the approach promises broad impact among the research community itself. Once the diagnostic assessment and reporting system is fully operationalized, other researchers can use it to study different aspects of instructional systems.

References

- Achieve (2015). *The role of learning progressions in competency-based pathways*. Washington, DC: Achieve, Inc. (July).
- Almond, R. G., Mislevy, R. J., Steinberg, L. S., Yan, D., & Williamson, D. M. (2015). *Bayesian Networks in Educational Assessment*. New York, NY: Springer.
- Ambrose, R., Baek, J. M., & Carpenter, T. P. (2003). *Children's invention of multidigit multiplication and division algorithms* (pp. 307-338). A. J. Baroody, & A. Dowker (Eds.). Mahwah, NJ: Taylor & Francis.
- Bell, A., Fischbein, E., & Greer, B. (1984). Choice of operation in verbal arithmetic problems: The effects of number size, problem structure and context. *Educational studies in Mathematics*, 15(2), 129-147.
- Beaton, A. E. (1996). *Mathematics Achievement in the Middle School Years*. IEA's Third International Mathematics and Science Study (TIMSS). Boston College, Center for the Study of Testing, Evaluation, and Educational Policy, Campion Hall 323, Chestnut Hill, MA 02167
- Black, P. (2004). *Working inside the black box: Assessment for learning in the classroom*. Granada Learning.
- Bolt, D. M., Ysseldyke, J., & Patterson, M. J. (2010). Students, teachers, and schools as sources of variability, integrity, and sustainability in implementing progress monitoring. *School Psychology Review*, 39(4), 612-630.
- Borko, H., Mayfield, V., Marion, S., Flexer, R., & Cumbo, K. (1997). Teachers' developing ideas and practices about mathematics performance assessment: Successes, stumbling blocks, and implications for professional development. *Teaching and Teacher Education*, 13(3), 259-278
- Briggs, D. C., & Peck, F. A. (2015). Using learning progressions to design vertical scales that support coherent inferences about student growth. *Measurement: Interdisciplinary Research and Perspectives*, 13(2), 75-99.
- Burris, C. C. (2014). *On the same track: How schools can join the 21st century struggle against resegregation*. Boston, MA: Beacon Press.
- Carlson, D., Geoffrey, D. B., Robinson, M. (2011). A multistate district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement. *Educational Evaluation and Policy Analysis*, 33(3), 378-398.
- Cai, L. (2010). A two-tier full-information item factor analysis model with applications. *Psychometrika*, 75, 581-612.
- Chazan, D., & Ball, D. (1999). Beyond being told not to tell. *For the learning of mathematics*, 2-10.
- Chung, G. K. W. K., O'Neil, H. F., & Herl, H. E. (1999). The use of computer-based collaborative knowledge mapping to measure team processes and team outcomes. *Computers in Human Behavior*, 15(3), 463-493.
- Clements, D. H. & Sarama, J. (2014). *Learning and teaching early math: The learning trajectories approach*. New York, NY: Routledge.
- Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher*, 32(1), 9-13.
- Confrey, J. (2012). Articulating a learning sciences foundation for learning trajectories in the CCSS-M. In Van Zoest, L. R., Lo, J.-J., & Kratky, J. L. (Eds.), *Proceedings of the 34th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*. Kalamazoo, MI: Western Michigan University.

Confrey, J. (2015). The value of learning maps and evidence-centered design of assessment to educational data mining. Briefing book for Computing Research Association workshop, "Advancing Data-Intensive Research in Education." Washington, DC, June 2.

Confrey, J., Hasse, E., Maloney, A. P., Nguyen, K. H., & Varela, S. (2011). A summary report from the conference designing technology-enabled diagnostic assessments for K-12 mathematics. *Designing technology-enabled diagnostic assessments for K-12 mathematics*. Raleigh, NC.

Confrey, J., Jones, R. S., & Gianopulos, G. (2015). Challenges in Modeling and Measuring Learning Trajectories. *Measurement: Interdisciplinary Research and Perspectives*, 13(2), 100-105.

Confrey, J., & Maloney, A. (2012). A next generation digital classroom assessment based on learning trajectories. *Digital teaching platform*, 134-152.

Confrey, J., Maloney, A. P., & Corley, A. K. (2014). Learning trajectories: A framework for connecting standards with curriculum. *ZDM*, 46(5), 719-733.

Confrey, J., Maloney, A., Nguyen, K., Mojica, G., & Myers, M. (2009). Equipartitioning/splitting as a foundation of rational number reasoning using learning trajectories. In M. Tzekaki, M. Kaldrimidou, & H. Sakonidis (Eds.), *Proceedings of the 33rd conference of the International Group for the Psychology of Mathematics Education* (Vol. 2, pp. 345–352). Thessaloniki, Greece: International Group for the Psychology of Mathematics Education.

Daro, P., Mosher, F.A., & Corcoran, T. (2011). *Learning Trajectories in Mathematics: A Foundation for Standards, Curriculum, Assessment, and Instruction*. Consortium for Policy Research in Education. www.cpre.org/learning-trajectories-mathematics-foundation-standards-curriculum-assessment-and-instruction

Delisle, J. R. (2015). Differentiation doesn't work. *Education week*, 34 (15). Retrieved January 8, 2015 from <http://www.edweek.org/ew/articles/2015/01/07/differentiation-doesnt-work.html?qs=differentiation>

Dougherty, B. (2008). Measure up: A quantitative view of early algebra. *Algebra in the early grades*, 389-412.

Dougherty, B. (2008). Measure up: A quantitative view of early algebra. *Algebra in the early grades*, 389-412.

Dweck, C. (2006). *Mindset: The new psychology of success*. Random House.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences*, 111(23), 8410-8415.

Greer, B. (1987). Nonconservation of multiplication and division involving decimals. *Journal for Research in Mathematics Education*, 37-45.

Greer, B. (1992). Multiplication and division as models of situations. De Corte, E., Greer, B., & Verschaffel, L. (1996). *Mathematics teaching and learning*.

Gummer, E. S., & Mandinach, E. (in press). Building a conceptual framework for data literacy. *Teachers College Record*.

Harel, G., & Confrey, J. (Eds.). (1994). *The development of multiplicative reasoning in the learning of mathematics*. SUNY Press.

Hansen, M. (2013). Hierarchical item response models for cognitive diagnosis. (Doctoral dissertation). Retrieved from ProQuest. (3567806) August, 2015.

- Heritage, M. (2008). *Learning progressions: Supporting instruction and formative assessment*. Washington, DC: The Council of Chief State School Officers.
- Heritage, M., & Popham, W. J. (2013). *Formative assessment in practice: A process of inquiry and action*. Cambridge, MA: Harvard Education Press.
- Hess, K. (2008). *Developing and Using Learning Progressions as a Schema for Measuring Progress*. National Center for Assessment. http://www.nciea.org/publications/CCSSO2_KH08.pdf
- Hess, K.K. & Kearns, J. (2011). *Learning Progressions Frameworks Designed for Use with the Common Core State Standards in Mathematics*. National Center for the Improvement of Educational Assessment. www.nciea.org/publication_PDFs/Math_LPF_KH11.pdf
- Hiebert, J., & Behr, M. (1988). *Number Concepts and Operations in the Middle Grades, Volume 2. Research Agenda for Mathematics Education*. National Council of Teachers of Mathematics, 1906 Association Drive, Reston, VA 22091.
- Izsák, A. (2008). Mathematical knowledge for teaching fraction multiplication. *Cognition and Instruction*, 26(1), 95-143.
- James, W. (1971). *William James: The essential writings* (p. xiii). Wilshire, B. W., Ed.). New York, NY: Harper & Row.
- Johnson, L., Adams Becker, S., Estrada, V. & Freeman, A. (2014). *NMC Horizon Report: 2014 K-12 Edition*. Austin, TX: New Media Consortium.
- Kim, M-J. & Lehrer, R. (in press). Using Learning Progressions to Design Instructional Trajectories. In C. Suurtamm (Ed.), *Annual Perspectives on Mathematics Education (APME) 2015: Using Assessment to Improve Instruction*. Reston, VA: National Council of Teachers of Mathematics.
- Kilpatrick, J. (2014). Competency frameworks in mathematics education. In Lerman, S. (Ed.) *Encyclopedia of mathematics education (85-87)*. Springer Netherlands.
- Lehrer, R., Kim, M.-J., Ayers, E., & Wilson, M. (2014). Toward establishing a learning progression to support the development of statistical reasoning. In Maloney, A. P., Confrey, J., & Nguyen, K. H. (Eds.) *Learning over time: Learning trajectories in mathematics education* (pp. 32-59). Charlotte, NC: Information Age Publishing.
- Lehrer, R. & Schauble, L. (2015). Learning progressions: The whole world is NOT a stage. *Science Education*, 99(3), 432-437.
- Mandinach, E. B., & Gummer, E. S. (2013). Defining data literacy: A report on a convening of experts. *Journal of Educational Research and Policy Studies*, 13(2), 6-28.
- Mehan, H. (2015). *Detracking: A Promising Strategy to Increase Social Mobility for Underserved Youth Opening the Doors to Opportunity for All: Setting a Research Agenda for the Future*, 29.
- Mislevy, R.J. (2006). Cognitive psychology and educational assessment. In Brennan, R. (Ed.), *Educational Measurement* (pp.257-305). Praeger Publishers, Westport, CT.
- Mosher, F. (2011). *The Role of Learning Progressions in Standards-Based Education Reform*. Consortium for Policy Research in Education. http://cpre.org/sites/default/files/policybrief/1218_lppolicybriefwebready.pdf
- Murphy, R., Gallagher, L., Krumm, A. E., Mislevy, J., & Hafter, A. (2014). *Research on the Use of Khan Academy in Schools: Research Brief*. Menlo Park: SRI International. Retrieved June 20, 2014 from http://www.sri.com/sites/default/files/publications/2014-03-07_implementation_briefing.pdf
- Northwest Evaluation Association. (2000). *Measures of academic progress*. Portland, OR: Author.

Oakes, J. (1990). *Multiplying inequalities: The effects of race, social class, and tracking on opportunities to learn mathematics and science*. Santa Monica, CA: The RAND Corporation.

Oakes, J. (2015). Research in the ongoing struggle for educational equity: An agenda for the future. In *Opening the doors to opportunity for all: Setting a research agenda for the future*. American Institutes for Research: Research Roundtable on Equity Issues in Education, January 2015 (pp. 8-16).

OECD (2012). Country note: United States. Programme for International Student Assessment (PISA) Results from PISA 2012. Retrieved from <http://www.oecd.org/pisa/keyfindings/PISA-2012-results-US.pdf>

Otto, A., Caldwell, J., Hancock, S. W., & Zbiek, R. M. (2011). *Developing Essential Understanding of Multiplication and Division for Teaching Mathematics in Grades 3-5*. National Council of Teachers of Mathematics. 1906 Association Drive, Reston, VA 20191-1502.

Patrick, S., & Sturgis, C. (2015). *Maximizing competency and blended learning: Insights from experts*. Vienna, VA: International Association for K-12 Online Learning. March.

Naumann, A., Hochweber, J., and Hartig, J. (2014). Modeling instructional sensitivity using a longitudinal multilevel differential item functioning approach. *Journal of Educational Measurement*, 51(4), 381–399.

Pane, J. F., Griffin, B. A., McCaffrey, D. F., & Karam, R. (2014). Effectiveness of Cognitive Tutor Algebra I at scale. *Educational Evaluation and Policy Analysis*, 36(2), 127 – 144.

Pane, J. F., Steiner, E. D., Baird, M. D., & Hamilton, L. S. (2015). *Continued progress: Promising evidence on personalized learning*. Santa Monica, CA: RAND Corporation.

Penuel, W. R., Fishman, B. J., Cheng, B. H., & Sabelli, N. (2011). Organizing research and development at the intersection of learning, implementation, and design. *Educational Researcher*, 40(7), 331-337.

Ready, D. (2014). *Student Mathematics Performance in the First Two Years of Teach to One: Math Teachers College, Columbia University*. http://www.newclassrooms.org/resources/Teach-to-One_Report_2013-14.pdf

Reise, S.P. (2012). The rediscovery of bifactor measurement models. *Multivariate Behavioral Research*, 47(5), 667-696.

Rupp, A., Templin, J., & Henson, R. (2010). *Diagnostic Measurement : Theory, Methods, and Applications*. New York, NY: Guilford Press.

Russell, S. J. (2012). CCSSM: Keeping teaching and learning strong. *Teaching Children's Mathematics*, 19(1), 50-56.

Steffe, L. (1994). Children's Multiplying Schemes.

Salas, E., Sims, D. E., & Burke, C. S. (2005). Is there a “Big Five” in teamwork?. *Small group research*, 36(5), 555-599.

Saxe, G. B., Shaughnessy, M. M., Shannon, A., Langer-Osuna, J. M., Chinn, R., & Gearhart, M. (2007). Learning about fractions as points on a number line. In W. G. Martin, M. E. Strutchens, & P. C. Elliott, (Eds.), *The Learning of Mathematics: 2007 Yearbook* (pp. 221–237). Reston, VA: NCTM.

Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist's perspective. *Journal for Research in Mathematics Education*. 26, 114-145.

Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, 10(4), 313-340.

Stephan, M. L. (2014). Conducting classroom design research with teachers. *ZDM*, 1-13.

Sztajn, P., Confrey, J., Wilson, P. H., & Edgington, C. (2012). Learning trajectory-based instruction toward a theory of teaching. *Educational Researcher*, 41(5), 147-156.

Tanenbaum, C., Le Floch, K., & Boyle, A. (2013). Are personalized learning environments the next wave of K-12 education reform. August. *American Institute for Research*. Retrieved from http://www.air.org/sites/default/files/AIR_Personalized_Learning_Issue_Paper_2013.pdf.

Tatsuoka, K. K. (2009). *Cognitive Assessment: An introduction to the Rule Space Method*. New York: Taylor & Francis Group, LLC.

Templin, J., & Bradshaw, L. (2014). The use and misuse of psychometric models. *Psychometrika*, 79(2), 347–354.

Vergnaud, G. (1994). Multiplicative conceptual field: What and why. *The development of multiplicative reasoning in the learning of mathematics*, 41-59.

Willms, J. D. (2003). *Monitoring school performance: A guide for educators*. Routledge.

Ysseldyke, J., & Tardrew, S. (2007). Use of a progress monitoring system to enable teachers to differentiate mathematics instruction. *Journal of Applied School Psychology*, 24 (1), 1-28. DOI: 10.1300/J370v24n01_01.

-(5(&21)5(<

-HUH&RQIUH
 -RVHSK 'LORRUH 'LVWLQJXLVKHG 8QLYHUVLW\ 3URIHVVURU
 67(0 'HSDUWPHQW
 &ROOHJH RI (GXFDWLRQ
 IRUWK &DUROLQD 6WDWH 8QLYHUVLW\
 5DOHLJK, 1& 27606

7HO: (919)513-8523
 (-PDLO: MFRQIUH#QFVX.HGX

352)(66,21\$/ 35(3\$5\$7,21:

'XNH 8QLYHUVLW\ &RUQHOO 8QLYHUVLW\ &RUQHOO 8QLYHUVLW\ ,XUKDP ,WKDFD ,WKDFD	0DWKHPDWLFV/3KLORVRSK\ 0DWKHPDWLFV (GXFDWLRQ 0DWKHPDWLFV (GXFDWLRQ	% . 6. 0.(G. 3K.!	1974 1979 1981
---	--	-------------------------	----------------------

\$332,170(176:

- 2007-SUHVHQQW *North Carolina State University*, -RVHSK 'LORRUH 'LVWLQJXLVKHG 8QLYHUVLW\ 3URIHVVURU RI 0DWKHPDWLFV (GXFDWLRQ, 'LUHFWRU RI 6FDOLQJ 8S 'LJLWDO 'HVLJQ 6WXGLHV (68"6)
- 2012-2014 *Amplify Learning, Durham, NC* &KLHI 0DWKHPDWLFV 2IILFHU (on leave)
- 2003- 2007 *Washington University*, 3URIHVVURU RI 0DWKHPDWLFV (GXFDWLRQ
- 1997-2003 *University of Texas at Austin*, 3URIHVVURU RI 0DWKHPDWLFV (GXFDWLRQ; 'LUHFWRU RI 6VWHPLF 5HVHDUFK &HQWHU; FR-IRXQGHU RI 8-7HDFK
- 1984-1997 *Cornell University*, \$VVRELDWH 3URIHVVURU RI 0DWKHPDWLFV (GXFDWLRQ
- 1981-1984 *Mount Holyoke College*, \$VVLVWQW 3URIHVVURU RI 0DWKHPDWLFV; 'LUHFWRU DQG)RXQGHU RI 6XPPHU0DWK, FR-IRXQGHU RI 6XPPHU0DWK IRU 7HDFKHUV
- 1980-1981 *Michigan State University*, \$VVLVWQW 3URIHVVURU RI 0DWKHPDWLFV

352'8&76/38%/,&\$7,216:

(L) 5HODWHG

&RQIUH\, -. & -RQH, 5. 6. (LQ SUHV). 8VLQJ OHDUQLQJ WUDMHFWRULHV WR SUHSDUH VL[WK JUDGH VWXGHQW LQ PHDVXUHPHQW DQG GDWD GLVSOD\ IRU LQYHQWLQJ D YDULDELOLW\ VWDWLVWLF. ,Q 3IDQQXFK, 0. & %HQ=-YL, ' . ((GV.) *The first international handbook of research in statistics education*, OHZ <RUN: 6SULQJHU.

&RQIUH\, -. , -RQH, 5. 6., & *LDQRSXORV, *. (2015). &KDOOHQJHV LQ ORGHOLQJ DQG OHDVXULQJ /HDUQLQJ 7UDMHFWRULHV. *Measurement: Interdisciplinary Research and Perspectives*, 13(2), 100-105. &RQIUH\, -. , & 0DORQH\, \$. 3. (2015). \$GHVLQJ UHVHDUFK VWXG\ RID FXUULFXOXP DQG GLDJQRVWLF DVVHVVPHQW V\VWHPL IRU D OHDUQLQJ WUDMHFWRU\ RQ HTXLSDUWLWLRQLQJ. *ZDM* 47(6), 919-932.

&RQIUH\, -. & 0DORQH\, \$. 3. (2014) /LQNLQJ VWDQGDUGV DQG OHDUQLQJ WUDMHFWRULHV: %RXQGDU\ REMHFVW DQG UHSUHVHQWDWLRQV. ,Q 0DORQH\, \$. 3., &RQIUH\, -. & 1JX\HQ, ..+. ((GV.), *Learning over time: Learning trajectories in mathematics education*. &KDUORWWH: ,QIRUPDWLRQ \$JH. (125-160).

6JWDMQ, 3., &RQIUH\, -. , :LOVRQ, 3. +., & (GJLQJWRQ, &. (2012). /HDUQLQJ WUDMHFWRU\ EDVHG LQVWUXFWLRQ: WRZDUG D WKHRU\ RI WHDFKLQJ. *Educational Researcher*, 41(5), 147±156.

(LL) 2WKHU 6LJQLILFDQW 3URGXFWV.

&RQIUH\, -. , 0DORQH\, \$. 3., & &RUOH, \$. .. (2014). /HDUQLQJ WUDMHFWRULHV: \$IUDPHZRUN IRU FRQQHFWLQJ VWDQGDUGV ZLWK FXUULFXOXP. *ZDM—The International Journal on Mathematics Education*. GRL: 10.1007/V11858-014-0598-7

0DORQH\, \$. 3., &RQIUH\, -. , 1J, ' . & 1LFNHOO, -. (2014) /HDUQLQJ WUDMHFWRULHV IRU LQWHUSUHVLQJ WKH -8

&&66-0: 0LGGOH JUDGHV VWDWLWLVF VWDQGDUGV DV DQ H[DPSOH. ,Q .DUS, . . . & 0F'XILH, \$. 5. ((GV.) *NCTM Annual perspectives in mathematics education 2014: Using research to improve instruction.* (23-34)

3HQXHO, -. 5., &RQIUH\, -, 0DORQH\, \$., & 5XSS, \$. \$. (2014). 'HVLJQ GHFLVLRQV LQ GHYHORSLOJ DVVHVVPHQWV RHOHDUQLQJ WUDMHFWRULHV: \$FDVH VWXG\'. *Journal of the Learning Sciences*, 23(1), 47±95.

&RQIUH\, -, & 0DORQH\, \$. 3. (2012). 1H[W JHQHUDLWLRQ GLJLWDO FODVVURRP DVVHVVPHQW EDVHG RQ OHDUQLQJ WUDMHFWRULHV LQ PDWKHPDWLFV. ,Q & 'HGH DQG - 5LFKDUGV ((GV), *Steps Toward a Digital Teaching Platform.* 1HZ<RUN: 7HDFKHUV & ROOHJH 3UHVV.

&RQIUH\, -, DQG 0DORQH\, \$. (LQ SUHVV). (QJLQHULQJ >RU@ HIIHFWLYHQHV LQ PDWKHPDWLFV HGXFDWLRQ: ,QWUHYHQWLRQ DW WKH LQVWUXFWLRQDO FRUH LQ DQ HUD RI FRPPRQ FRUH VWDQGDUGV. ,Q -LQID & DL DQG -DPHV 0LGGOHWRQ ((GV). *Large Scale Studies in Mathematics.* 6SULQJHU.

6<1(5*,67,& \$&7,9,7,(6:

- x *Bill and Melinda Gates Foundation* 'HVLJQ RI D OHDUQLQJ PDS IRU PLGGOH VFKRRO PDWKHPDWLFV
- x *Winning team:* 7ULDQJOH 6WDUWXS :HHNHQG, 5DOHLJK, 1&, 6ROYLQJ SUREOHPV LQ WKH -.12 \VWHP, 2015
- x *North Carolina State University Innovator of the Year Award,* 2012
- x *Validation Committee:* \$ PHPEHU RI WKH QDWLRQDO YDOLGDWLRQ FRPPLWWHH IRU WKH *Voluntary State Common Core Standards* VSRQVRUHG E\ WKH IDWLRQDO *RYHUQRU\ VVRFLDWLRQ DQG &RXQFLO RI &KLHI 6WDWH 6FKRRO 2IILFHUV, 2011
- x *NAEP Committee on The Future Of NAEP,* 2011
- x *Chair of the National Research Council Committee for a Review of the Evaluation Data on the Effectiveness of NSF-Supported and Commercially-Generated Mathematics Curriculum Materials* ± WKH SURMHFW UHVXOWHG LQ WKH SXEOLFDFWLRQ RI "2Q (YDOXDWLRQ & XUULFXODU (HIIHFWLYHQHV: -XGJLQJ WKH 4XDOLW\ RI -.12 0DWKHPDWLFV (YDOXDWLRQV".
- x *Co-Founder of the UTeach program at the University of Texas in Austin*

&2/(\$%25\$7256 & 27+(5\$),/, \$7,216:

(L) &ROODERUDWRUV DQG &R-(GLWRUV 'XULQJ WKH 3DVW)RXU <HDUV - 727\$/ 9
&KULV 'HGH, +DUYDUG 8QLYHUVLW\ 5REHUWV 5H\, 8QLYHUVLW\ RI 0LVVVRXUL
5LFKDUG /HKUHU, 9DQGHUELOW 8QLYHUVLW\ \$QGUH 5XSS, (76
:LOOLDP 3HQXHO, 8QLYHUVLW\ RI &RORUDGR 3DROD6]WDMQ, IRUWK & DUROLQD 6WDWH 8QLYHUVLW\
%DUEDED 5H\, 8QLYHUVLW\ RI 0LVVVRXUL /XWKHU :LOOLDPV, 7XVNHJHH 8QLYHUVLW\
.RHQR *UDYHPHLMHU, 7HFKQLVFKH 8QLYHUVLW\ (LQKRYHQ +ROW :LOVRQ, 81& *UHHQVERUR
6XVDQQH 3UHGLJHU, 7HFKQLVFKH 8QLYHUVLW\ 'RUWXPQG

(LL) ,QYHVWLJDWRU\ *UDGXDWH DQG 3RVWGRFWRUDO \$GYLVRUV - 727\$/ 2
3K'. \$GYLVRU: +DUULVRQ *HLVHOPDQ, &RUQHOO 8QLYHUVLW\; .HQQHWK 6WULNH, &RUQHOO 8QLYHUVLW\
0.6. \$GYLVRU: 'DYLQ +HQGHUVRQ, &RUQHOO 8QLYHUVLW\

(LLL) 7KHVLV \$GYLVRU DQG 3RVWJUDGXDWH-6FKRODU 6SRQVRU- 727\$/ 24
3K'. 6WXGHQWV: (7RWDQ 24);
Completed: \$QGUHZ &RUOH\, .HQQ\ 1JX\HQ (1& 6WDWH), (ULQ .UXSD (0RQWFODLU 6WDWH 8QLYHUVLW\), 6KD\0D
7KRPDV (.LQNDLQ 6FKRRO), *HPPD 0RMLFD (81& &KDSHO +LOO), +ROW = LOVRQ (81& *UHHQVERUR), 6LEHO
.D]DN (7XUNH\), .DWLH 0DNDU (8 RI 4XHHQVODQG), -HQQLIHU :LOKHOP (8 RI .HQW\FN\), (OL]DEHWK 5RZH
(G*(DW 7(5&), 0DULD 0HSHWLX ((XURSHDQ 8QLYHUVLW\, &SUXV), -RVH \$LUHV & DVWUR-)LOKR ()HGHUO
8QLYHUVLW\ RI &HDUI), 6KDZQ +DDUHU, \$QGUHD 0DULH /D&KDQFH (681<, &RUWODQG), 'DYLQ 'HQQLV, +HQQ
'RHUU (6UDFXVH 8QLYHUVLW\), 6XVDQ 3LOLHUR (&RUQHOO 8QLYHUVLW\), (ULFN 6PLWK, 0DUFHOR %RUED (6DR 3DROR
6WDWH 8QLYHUVLW\), .DUROLQH \$IDPDVDJD-)XDWD'L, -DQ 5L]XWL, :HQG\/. 0LOOUR\

In Progress: %ODNH :LWQH\ (1& 6WDWH), 0HHWDO 6KDK (1& 6WDWH)