CADRE REPORT

Teacher Perspectives on the Content-Referenced Growth Reporting Prototype: Findings from Interviews

By Derek Briggs, Olivia Cox, Sandy Student, Erik Whitfield*

June 2023

Prepared by the Center for Assessment, Design, Research and Evaluation (CADRE) at the CU Boulder School of Education.

**Authors listed in alphabetical order, all contributed equally to the writing of this report.*

School of Education
UNIVERSITY OF COLORADO **BOULDER**

Abstract

This study aims to analyze in-service elementary teachers' reactions to a prototype for reporting content referenced growth on a widely used commercial assessment, Curriculum Associates' i-Ready Diagnostic. The prototype is built around an approach to growth that centers learning progressions (LPs), tools that describe the developmental path students are likely to take when learning a big picture concept in mathematics or reading (Clements & Sarama, 2004). Created through a partnership between CU Boulder and Curriculum Associates, the score reporting prototype is intended to support teachers, parents, and students in interpreting both the status of student understanding at one point in time and their growth in understanding across points in time. The purpose of this study is to gain insight into the intended use of the prototype by soliciting feedback from practicing teachers about: (a) the potential usefulness of embedding LP information into the i-Ready diagnostic report, (b), substantive interpretations supported by the LP prototype, and (c) the prototype's usability.

Introduction

The objective of this qualitative study was to analyze in-service elementary educators' reactions to a prototype for reporting content-referenced growth (CRG) on Curriculum Associates' i-Ready Diagnostic. The prototype, accessible with this link, is built around an approach to conceptualizing growth that centers learning progressions (LPs), which are conjectures about the developmental path students are likely to take as they receive instruction related to a big-picture concept in mathematics, science or reading (e.g., Clements & Sarama, 2004). As long as it is possible to generate a reasonable hypothesis of development, taking a LP approach to growth is feasible, although it also requires adequate theoretical and empirical support. In the context of the CRG project, this consists of using prior research to define an LP,

then investigating the extent to which the LP is predictive of item difficulty ordering, as described in Wellberg, Briggs and Student (2022).

We designed the prototype around an LP that describes how children develop increasingly sophisticated understandings of fractions (Wellberg, Briggs, & Student, 2022). The prototype itself is intended to support teachers, parents, and students in interpreting both the status of student understanding about fractions at a single point in time, and their growth in understanding across points in time. In this study, we solicited feedback from practicing teachers who are also i-Ready users to learn more about: (a) the potential usefulness of embedding information from the fractions LP into existing i-Ready diagnostic score reports, (b) ways in which the LP prototype might support substantive interpretations about both students' scale scores and conceptual understanding of fractions, and (c) the prototype's general usability.

Specific research questions included:

1. To what extent can an LP approach to growth, as represented by the prototype, facilitate meaningful interpretations about differences in student scale scores on the i-Ready Diagnostic, and how?
2. To what extent are teachers able to understand and interpret the LP by using the prototype, and how?
3. How do teachers envision using information provided by the prototype?

Methods

In summer 2022, our team developed a protocol for interviewing teachers after some initial pilot interviews with former classroom teachers, now graduate students at CU Boulder. We designed the protocol to include open-ended questions with minimal pre-teaching about the prototype. This approach made it possible to learn how teachers made sense of what they were seeing on their own without any scene-setting from us.

The final version of the protocol is divided into three sections: (a) an introduction in which we introduced ourselves and provided a brief overview about the purpose and goals for the interview, (b) a feedback session in which we invited teachers to explore the prototype and then asked them a series of five questions, and (c) a final wrap up in which we posed more general questions, reviewed logistics, and thanked teachers for their time. The final version of the protocol, which is included as Appendix A of this report, reflected our efforts, with guidance from Curriculum Associates' UX Team, to make questions as open-ended as possible to avoid constraining teachers' interpretations of the prototype. For example, early drafts of the protocol assumed parent-teacher conferences were a use case for the prototype and directed teachers to a particular view in the prototype, asking questions to observe the information they would extract. In the final version, we removed that section and instead asked, "In your instructional planning or practice, how might you use what you've learned from this class-level view, if it all?" The updates we made to the protocol through cycles of collaboration with our UX partners were crucial to the success of the interviews and facilitated valuable learning and reflection for our team as we moved through various stages of this study.

Sample

Recruitment for this study began with pre-screening using a short survey sent out to a large group of educators by one of our collaborators at Curriculum Associates. Information about educators collected on the survey included:

- Geographic location of educator's school district

- Grades taught in 2020-2021
- Years of experience as teacher of record
- Years of experience using i-Ready Diagnostic
- Demographic information about the educator (gender, race/ethnicity)
- Demographic information about the students in school district (race/ethnicity, FRL eligibility, and English language learner status)

A small group of 8 educators who expressed interest in participating were selected for us by Curriculum Associates according to the agreed upon criteria of ensuring some degree of representation across grades within the 3-6 range, geographic locale, and level of experience as a teacher of record. One of the educators was unable to make their scheduled interview time. Table 1 below contains information for all seven educators in the final sample. Note that four of the seven individuals were former classroom teachers who were presently serving as instructional coaches. We included classroom teachers and instructional coaches in our sample because individuals in both roles frequently use data to drive instruction (Snyder and Delgado, 2019). (All names shown in Table 1 are pseudonyms.)

Table 1. *Information on Teachers in Sample*

Name	State	Grade	Role
Danielle	MO	3, 4	Classroom teacher
Natalie	CA	5	Instructional coach
Rachel	NY	5	Classroom teacher
Geneva	CA	5	Classroom teacher
Andrew	CA	3, 5	Instructional coach
Courtney	IO	3, 4, 5	Instructional coach
Catherine	MA	4, 5	Instructional coach

We conducted all seven interviews over Zoom and they lasted between 40 and 60 minutes. In each interview, anywhere from two to four members of the research team were present. One person led each interview, while other team members took notes, speaking only occasionally as required to clarify specific points. All participants gave consent for the interviews to be recorded. As a result, all interviews save one come with the artifact of a video recording and transcript; recording difficulties forced us to record one interview by audio only. During the interviews, we loaded the reporting prototype on one of our computers and then gave the educator “mouse control” so that they could interact with the prototype directly without ever sharing their own screen.

Coding methods

For this analysis, we conducted two separate coding exercises to settle on our findings across the seven interviews. We used both deductive and inductive analysis to generate the themes outlined in the Results section below.

Deductive coding is a “top-down” approach to analysis in which a researcher starts with a set of guiding questions or themes, then looks for evidence within the interviews corresponding to these themes specifically (Maxwell, 2012). For the deductive portion of our analysis, we began by identifying four high-level themes as a research team. These themes represented organizational “bins” into which data could be sorted (Maxwell, 2012), and they pre-existed data collection because they consisted of categories, ideas, and concepts we predicted might arise based on our interest in how the prototype might facilitate meaningful interpretations about differences in i-Ready Diagnostic scale scores (RQ 1), as well as how understandable the prototype is to teachers (RQ 2) and how useful teachers perceive the prototype to be (RQ 3). To complete the deductive cycle of coding, one researcher reviewed transcripts and video recordings for instances in which these anticipated themes came up, transcribing

quotes ranging from one to several sentences as evidence of interviewees' attitudes and perceptions on each theme. The research team then reviewed these findings together.

While our deductive process began with agreed-upon themes we expected to be salient, we also engaged an inductive coding approach, which uses inductive logic to “infer categories or conclusions based on data” (Kennedy and Thornberg, 2018). Inductive analysis thus refers to a “bottom-up” approach in which researchers review transcripts to identify common points without any prior guidance on what to look for. That is, themes from inductive analysis are emergent: they represent what we heard without regard for what we were specifically looking for when we conducted these interviews. Completing an inductive coding cycle along with the first deductive cycle proved particularly useful for gaining insight into each of our RQs and turning our attention toward what emerged as important in the data with no expectation about what those important aspects could or should be. For this report, our procedure for inductive coding was to split the interviews into two sets alternating by the order in which the interviews were conducted. Using video recordings and text transcripts, two researchers from our team then reviewed the interviews–one interviewer per set of three/four. Each researcher identified their own themes and transcribed quotes ranging from one sentence to about one paragraph as evidence for these themes’ prevalence in the interviews. Once this was complete, these two researchers met to identify commonalities among their themes. This produced a final set of five themes, presented in the Results section.

Results

Deductive Coding

In this section, we present findings related to deductive codes for topics that guided the questions we asked in our interviews with teachers. First, we present evidence relative to four themes: (a) how teachers made connections between i-Ready scale scores and levels of the fractions LP depicted in the prototype, (b) how the prototype supported inferences about student growth and progress, (c) how teachers responded to data as it related to grade-level performance, and (d) teachers' interpretations of score uncertainty. Below, in Table 2, we note the

number of interviews–up to seven–in which we found evidence related to each of these themes. We break out evidence into strongly and weakly positive or negative.

Positive evidence means that we documented clear evidence of the prototype supporting the teacher relative to the theme in question, while negative evidence indicates that there was room for improvement. Moreover, we interpreted statements as strong versus weak types of evidence depending on how quickly or directly a teacher made any interpretation related to a theme. If, for example, a teacher interacted with a feature of the prototype and immediately made a statement inferring something about student growth, we flagged it as strong evidence of theme (b). If, however, the teacher required prompting to make a similar inference or omitted some parts of the desired inference, we considered their response an instance of weak evidence. Table 2 outlines the frequency and strength and nature of evidence relative to each of the first four themes. We also describe a fifth theme, (e) suggestions for improving the prototype. Because these suggestions often touched on (a) and (b), we documented these suggestions relative to both themes and did not count them up separately.

Table 2. *Count of interviews in which each Deductive Theme was Identified*

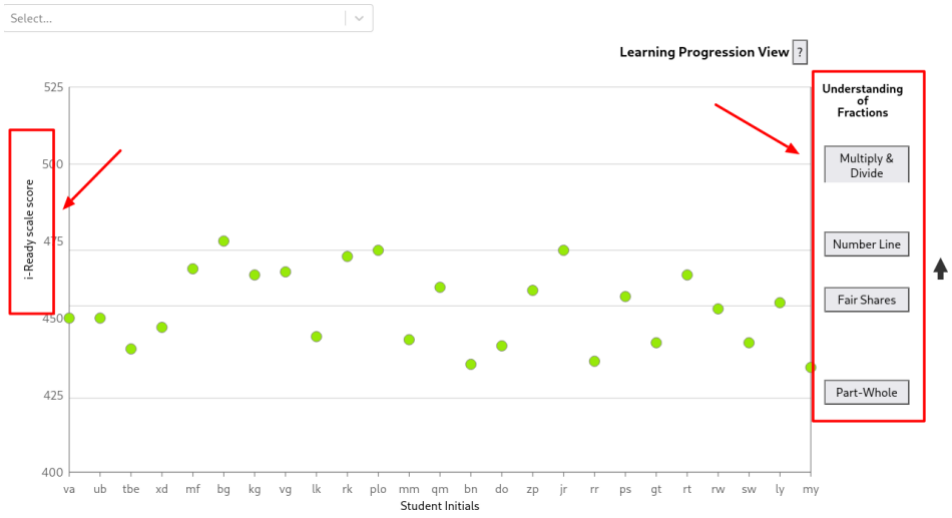
	Positive evidence		Negative evidence	
Theme	Strong	Weak	Strong	Weak
Connections between i-Ready Scale Scores and LP Levels	7	0	0	2
Inferences about student growth and progress	6	1	1	2
Using the prototype to make inferences about grade level expectations	3	2	0	4

Teachers' interpretations of score uncertainty	2	2	0	1

Connections between i-Ready Scale Scores and LP Levels

The first theme, connections between the i-Ready scale and LP levels, often came up when teachers dragged their mouse from the scale score on the left side of the display to the LP level labels on the right side (see Figure 1). Every teacher made a direct connection between the scale and the LP, as indicated by Table 2.

Figure 1. *Scale scores and LP level labels*



This theme involved teachers noticing and interpreting how any given i-Ready scale score might correspond to a different level of student understanding of fractions. For example, when asked to describe what he was seeing and what it might represent, Andrew began to connect scale scores and LP levels by first reading the names of each level and then saying, “I'm guessing that those correspond to the scale score, and the types of problems that they'll be getting, that they either would get or they would need to understand.” Then, turning his attention to the green dot representing one student, he went on, “so if I’m looking at LY… It's just above 450,

455. They understand fair shares, and their next progression would be to go towards Number Line understanding."

Similarly, upon navigating to the student-level view, Catherine moved her mouse from the left side of the y axis, which displayed the i-Ready scale score to the right side, which contained the LP levels, and said, "let's see, they're up to 475, so they're at the level where they can do things related to the number line." Andrew and Catherine were not the only teachers to make these kinds of remarks. According to our analysis, all seven teachers made statements that we interpret as relatively strong evidence in favor of the idea that the prototype supports users in drawing connections between the i-Ready scale and LP levels (see Table 2).

Two teachers also made comments that we interpreted as relatively weak evidence of ways in which the prototype might lead to inaccurate interpretations about the scale-LP relationship. For instance, Geneva equated a student's location on the scale to "mastery" of the associated LP level, though she quickly grasped that the scale location corresponds to a 50% likelihood of a correct response once it was explained. Rachel also made a comment that the LP levels appear "close" to one another on the scale, which our research team interpreted as a possible indication that the presence of the LP affected how this potential user made sense of the distance between the lower and upper portions of the LP. In reality these span significant portions of several grades' score distributions.

Inferences about student growth and progress

Teachers made many statements that were indicative of how the prototype might help them make inferences about student growth and progress. All seven teachers spoke favorably in this regard, and six of them delivered remarks that strongly suggest that the prototype supported their reasoning about changes in

students' proficiency (see Table 2).

We were interested in learning how teachers described changes in student understanding over time (growth) and how teachers might envision instructional next steps for the student based on their current conceptualization of fractions (progress). Statements about growth and progress most frequently occurred while teachers were in the "student view" of the prototype (see Figure 2) and in the "classroom view" when teachers clicked the checkbox to show prior scores (see Figure 3).

Figure 2. *Individual Student View*

Figure 3. *Classroom view with prior scores*

A number of teachers made sense of student growth by considering the three testing occasions depicted in the prototype by color-coded dots. For instance, when looking at the view that includes prior scores (Figure 3), Catherine remarked, "this shows growth. This is what we're really interested in. They have a fall, a winter, and a spring." She then clicked on each of the dots representing the testing occasion and said, "you can click on it, it's 411 to 435. So, they made more than a year's growth in a year. Because typical growth is anywhere from 12 to 15 points, which is what we've seen." In this way, Catherine used the prototype to make inferences about student growth, while drawing on her pre-existing conceptualizations of typical growth as an i-Ready user in her school context[1].

[1] This link elaborates on how the i-Ready Diagnostic defines Typical Growth, Stretch Growth, and other growth measures

While Catherine alluded to scale score points in her descriptions of growth when viewing prior scores, Andrew spoke about growth in terms of how students had progressed in levels of understanding on the LP, explaining, “they have a decent understanding of how it [fractions] can be represented on a number line. They are seeing the relationships between the fractions and decimals… they're not really ready for the multiplication and division.” Throughout interviews like Andrew’s and Catherine’s, we saw examples of teachers using the graphical information from the prototype to make inferences about how much students had grown over time and potential next steps to continue instructional progress.

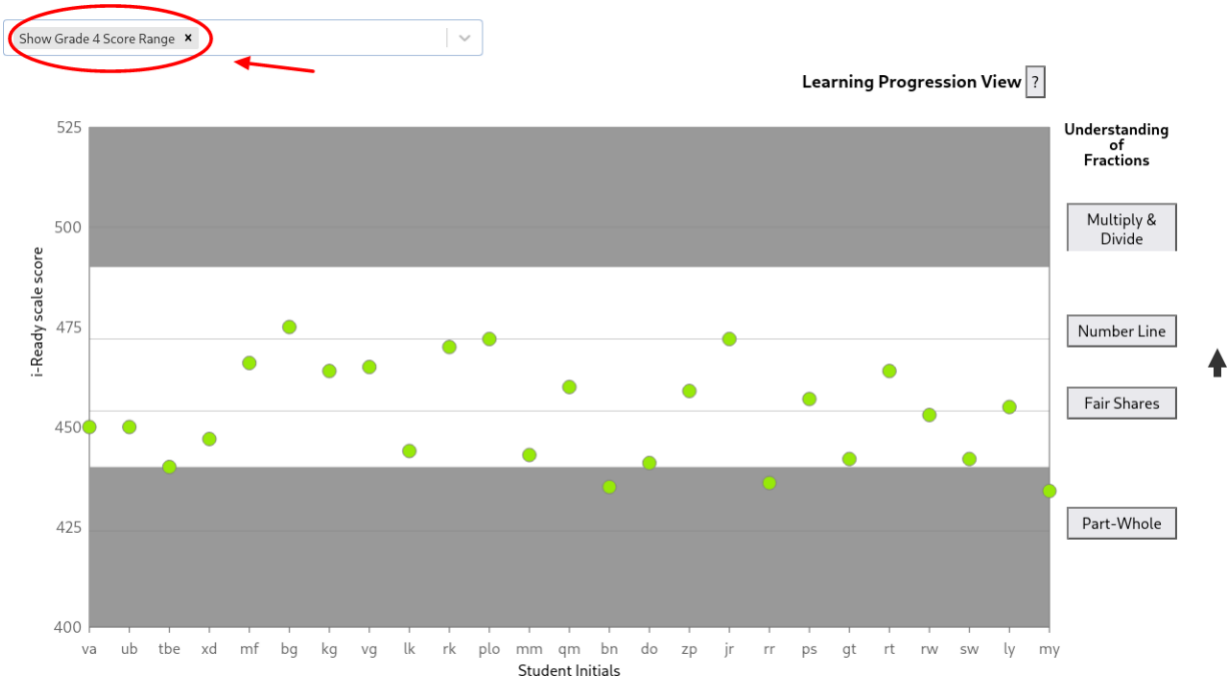
The evidence we tagged as negative was somewhat varied. The only strong evidence we found of misinterpretation about growth was Rachel’s comment that “I'm gonna assume that some of them are diagnostic scores and some of them are growth monitoring scores[2].” Otherwise, teachers appeared to quickly grasp that the scores came from the Diagnostic. As we touch upon below, we also received feedback (weaker negative evidence on theme 2) about the use of a single exemplar item.

[2] i-Ready offers the i-Ready Diagnostic assessment which is given 3 times a year (Fall, Winter, Spring) and the Growth Monitoring assessment which is a shorter assessment that can be given to students in months in which they don't take the i-Ready Diagnostic. Rachel’s statement indicates that she may have experience with both types of assessment at her school.

Using the prototype to make inferences about grade level expectations

The third theme, “grade level focus,” refers to moments in which teachers were concerned with the relationship between the content that appeared in the prototype and grade level proficiency bands or benchmarks for fourth grade. To evoke this focus, teachers tended to use the phrases, “on grade level” or “below grade level,” particularly in the “Show Grade 4 Score Range view” (see Figure 4). Most teachers made positive comments regarding the way the prototype surfaces information about student’s grade levels–mainly the inclusion of visual cues for “grade level” in the class- and student-level views–but we also received substantial feedback on how it could do so more effectively and in ways that better align with the realities of elementary education.

Figure 4. *Classroom view with score range*



As shown in Table 2, along with the five teachers who made positive comments related to the idea of grade level information, four teachers also delivered critical feedback that highlighted paths to improvement. Three teachers suggested providing clearer information about students' grade levels, whether they are "on grade level" or not. For instance, Catherine told us, "When I put my cursor over the dot, it tells me the student, the actual score, when they took it… doesn't tell me the grade level. But the teacher would know the grade level." Adding a student's grade level according to their diagnostic score was recommended here regarding the tooltip that appears when hovering over a student's scale location, as well as regarding the idea of a "grade 4 score range." Natalie suggested score ranges for grades 3 and 2 as well to better capture the same level of information across classes with students below grade level.

Teachers' interpretations of score uncertainty

In another deductive finding, we had anticipated that teachers could have trouble making sense of the view in the prototype that shows score uncertainty. Clicking this view converts single green dots to oval shapes that represent the range of scale scores each student might have received if they had taken a different subset of items (see Figure 5).

Figure 5. *Classroom view with score uncertainty*

We worried that the change from dot to oval might confuse teachers. Instead, four teachers interpreted the score uncertainty view with little or no explanation from us. For example, when Geneva first clicked on "Show Score Uncertainty," she wondered if the dots representing

students became blurry because she'd done something wrong, but quickly followed up with her own interpretation,

> is this maybe like, you know how it's blurry under and then above, is this where that student would really fall if maybe there was a change in how they answered the question, if they got lucky a few times or maybe they got a question wrong, kind of the range of where they could truly be?

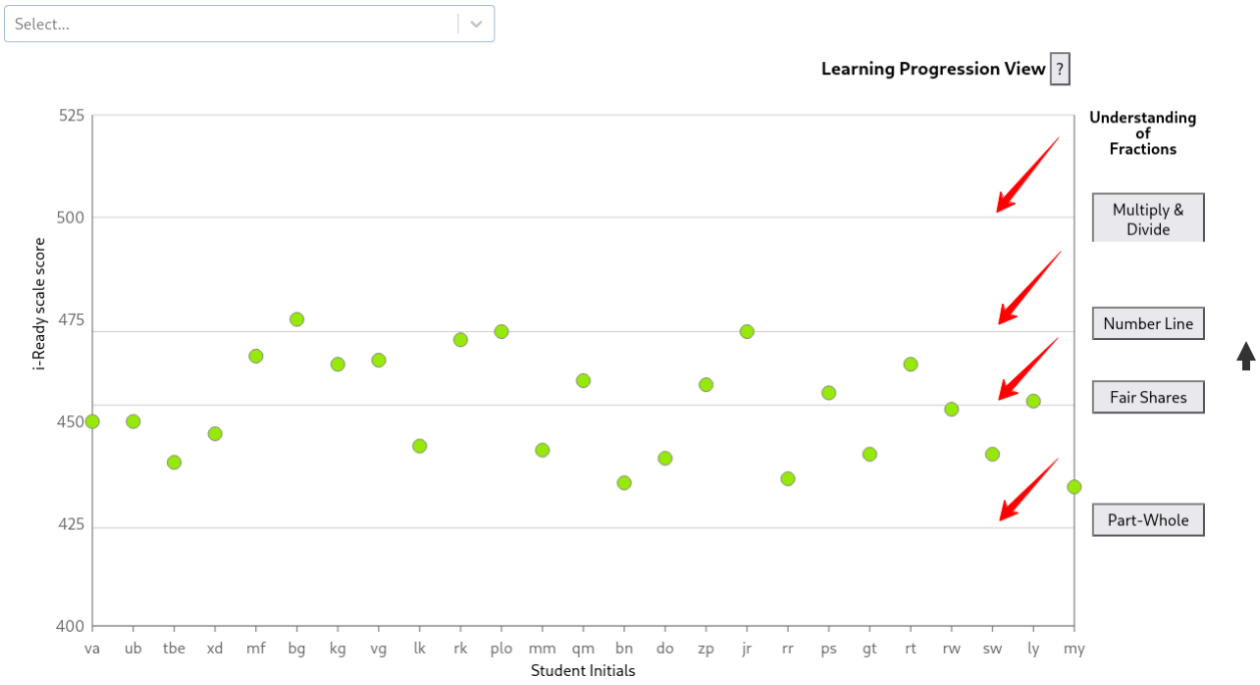
Andrew also showed some familiarity with the concept of score uncertainty, remarking that "a margin of error is included" in the view. These and other reactions provided us with some preliminary reason to believe that teachers may be more comfortable with a score uncertainty range than we initially assumed, especially since teachers made conclusions about what they were seeing and what it meant without much context-setting from us. Only one teacher indicated struggling with this feature; Catherine told us, "I'm not sure what show score uncertainty means. What are we trying to get at with that." After Olivia explained our intentions to Catherine, she responded, "I see. I wouldn't have understood that at all unless you explain it to me. Teachers would need to know what they're looking at." Hence, although it is not a given that all teachers will intuit the correct meaning of the "score uncertainty" feature, it is also not to be assumed that teachers could not make sense of this information when provided with appropriate support.

Suggestions for improving the prototype: interpretations of LP level thresholds and confounds in choice of exemplar items

We explicitly asked teachers about ways to improve the prototype and received feedback on possible improvements throughout each interview. Teachers mainly provided

substantive feedback around two recurring topics corresponding to deductive themes documented above. The first topic involved interpreting thresholds for different levels of the LP, and it corresponds to deductive theme (a), the connection between the scale and LP. The second topic involved challenges in the use of exemplar items to illustrate conceptual distinctions in LP levels, and it corresponds to deductive theme (a) as well as theme (b), inferences about growth and progress. Regarding the first topic, the LP thresholds refer to the horizontal lines that connect the i-Ready scale score on the left side of the display to each LP level on the right side (see Figure 6).

Figure 6. *Default classroom LP view with threshold focus*



These lines indicate the scale score a student would need to earn to have a 50% chance of correctly answering an item of average difficulty from each LP level. For example, the horizontal line connecting the Fair Shares level to the i-Ready scores is around

451, which means that students with an i-Ready Diagnostic scale score of 451 have a 50% chance of correctly answering a Fair Shares item of average difficulty.

In the interviews, teachers had a variety of reactions to this method of setting the LP level thresholds. None of the teachers arrived at the desired interpretation of the horizontal lines, so we decided to share our process for setting the thresholds at 50% and ask for their opinion. When Sandy, the interview facilitator, explained the process and asked Geneva for her thoughts on whether the 50% threshold seemed intuitive, she said,

> I would think it'd be higher…just because I always base it on 70% with teaching.
>
> Sandy: So, you think it would be informative to have it show mastery?
>
> Geneva: Yeah.

Andrew initially referenced the students clustered around the horizontal line connecting a 475 scale score to the Number Line LP level and explained how he would use the description of the next level above, Multiply and Divide, to identify what to teach or what was " keeping [students] back from the next level." However, when Olivia elaborated on the 50% threshold and asked if that information would change Andrew's previous interpretation, he stated, "Yeah, because that tells me all right. They have not mastered number line. They have a coin flip chance of understanding number line."

Like Andrew and Geneva, Danielle's first interpretation of the horizontal line corresponding to the Number Lines level was that it "means they've just about got that, so we're ready to move on." However, once we talked to her about the process for setting each threshold, she explained that it would not pose a problem for her if she knew what the level demarcations were ahead of time. Taken together, these comments from teachers suggest two considerations for improving the prototype: first, that we might rethink the 50% threshold for

setting LP levels given teachers' potential preference for figures that represent mastery, and second, that we ensure that the future thresholds, no matter where we set them, are transparent to users.

In another topic involving improvements to the prototype, three teachers commented on the difficulty of the Fair Shares level, stating that it might be too hard or contain too many concepts. Teachers often shared these comments after the portion of the Interview Protocol when we invited them to click through the LP level descriptions and example items. Danielle described the distance from the Part-Whole to the Fair Shares levels as "a big jump." She went on to explain that she would "almost want to have a separate addition and subtraction of fractions." Natalie made a similar observation, noting that she "might expect them to get [Number Line] before they get [Fair Shares]." Catherine was unambiguous in her assessment of the Fair Shares level, stating that "Fair Shares is more difficult than Number Line." To us, these reactions signaled something to keep in mind in future iterations of the prototype: that the characteristics of exemplar items are particularly important to help users conceptually distinguish between LP levels.

Inductive Coding

Table 3 presents counts for positive and negative evidence, weak and strong, related to each of five inductive themes: ease of use, use for small group instruction, color coding, communication to parents and students, and teacher professional development. Because the feedback we received on small group instruction and color coding overlapped substantially, we summarize evidence on both these themes in the section "Potential use of prototype for small group instruction."

***Table 3*. Count of Interviews in which each Inductive Theme Emerged**

Theme	Positive evidence		Negative evidence	
	Strong	Weak	Strong	Weak
Ease of use	2	3	0	0
Use for small group instruction	0	6	0	0
Color coding	1	3	0	2
Communication	2	1	0	0
Teacher PD	2	3	0	0

Ease of Use

An emergent theme was the extent to which teachers perceived the prototype to be easy to use and whether it would be easy for other teachers to use without additional training. As teachers navigated the drop-down menu to manipulate the presentation of data, five of the teachers remarked on the ease with which they could find relevant information (indicated in Table 3), including individual and group scores, previous scores, and students ordered by score. Specifically, they noted that information was accessible without tedious navigation and described ways in which the prototype might contribute something new to i-Ready's existing

score reporting system. For example, upon toggling to the class-level view, Rachel compared the prototype to current i-Ready score reporting tools, stating,

> It's nice to see the grouping for the kids and where you could put them . . . when I look at the Diagnostic for i-Ready, it's usually more individualized for students. I like the class view…Once I understood how to use the tool, I think I do like the data you have on there. It's an easier tool for me to read.

Reactions like Rachel's indicated that teachers found the prototype to be easily interpretable and that it represented a salient complement to the kind of score reporting experience already offered to i-Ready users.

Teachers also highlighted the value of having all the information necessary for a given instructional task, such as grouping students or interpreting standards, in one location. For instance, Courtney noticed that "on one screen" she could "find out a lot of information without having to jump back and forth" to various multi-page reports. Similarly, Danielle appreciated "how you can see all that data in one spot." Andrew also had a positive reaction to the idea of having class-level and student-level data at his fingertips, explaining,

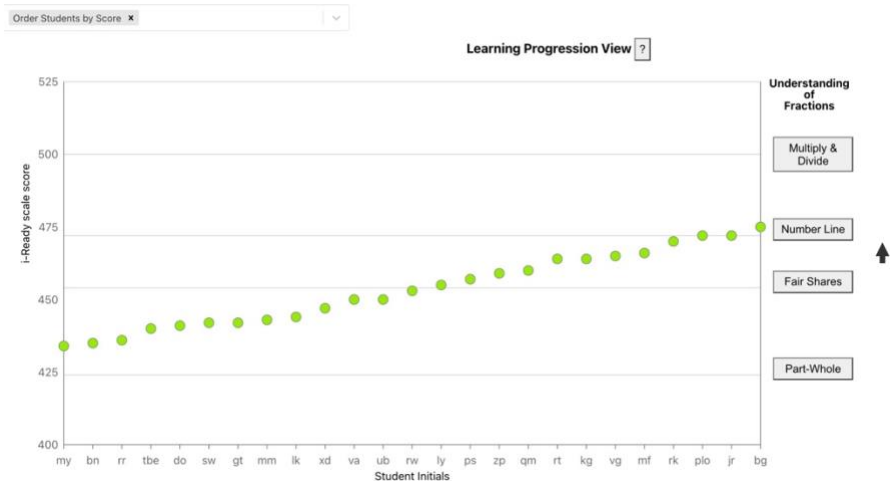
> "this (mouses over level labels), it's a little bit easier to read the learning progression, you know, because I know you can run like a learning progression report (in existing score reporting tools), but this is more like (closes fist and open fingers in explosion gesture), and there it is.

Overall, the teachers were excited by the prototype's capacity to facilitate speedy retrieval of information, including data about individual students and class-wide trends.

Potential utility of prototype for small group instruction

A second emergent theme was a desire to divide a whole class into small groups for intervention or differentiated instruction and how the prototype might support the creation of small groups. All but one teacher made at least one favorable comment about the usefulness of the prototype for grouping tasks, though we did classify all evidence as "weak" given the relative lack of detail in these brief comments (see Table 3). To this end, the teachers made frequent use of the sorting function in the drop-down menu, which ordered all the students in the hypothetical fourth grade class by their latest score on the i-Ready Diagnostic (see Figure 6).

Figure 6. *Classroom View with Students Ordered by Score*



Using her mouse to draw circles around a few green dots at a time, Rachel said, "I would probably group . . . like the ones that are close together . . . Now you can immediately see which ones I need to group together." When she clicked the option to order students by score, Courtney reacted by saying, "oh, and then this would order the students. This would help, I guess, with cluster grouping or small group instruction as well." Catherine also noted the usefulness of this view for small groups, noting that it made it easier to distinguish the needs of

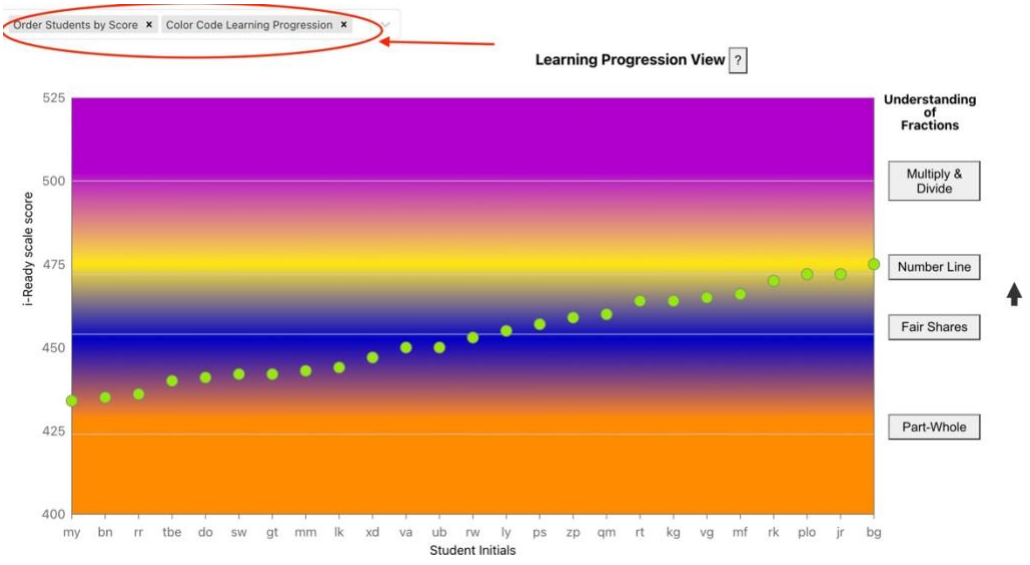
a “diverse group of learners.” Danielle immediately related what she saw in the prototype to her daily instructional practices.

> For grouping for interventions this would be wonderful . . . we dedicate a half hour every day to one specific intervention time, and so to be able to see it like this and say 'here's your progression that we want them to do in fractions, here's where this group of kids is,' just to see it like this I feel like would help teachers in their groupings

With statements like these, teachers described ways in which they might leverage information about students from the prototype to structure small group instruction and place certain students in certain groups.

An additional feature of the prototype that proved relevant for this goal of grouping students was the option for color coding the LP, which converted the previously white background of the class view to a set of multicolored horizontal bands (see Figure 7). Four teachers had positive comments about this feature (see Table 3).

Figure 7. Class view, color coded LP with students ordered by score



Geneva found that utilizing the color code feature while sorting students by score helped her "group them a little bit better," and that she would "try to band them with . . . scores that are similar within the same color range." The color-coded LP not only supported Geneva in identifying students who may be grouped together, it also allowed her to make conclusions about each group's current level of understanding of fractions and what they should be working toward. When viewing color coded LP with students ordered by score, she drew a circle with the mouse around a cluster of dots in the purple band representing the Fair Shares level, right underneath the yellow band representing the Number Line Level. Geneva then said, "okay so this is something I'd really like to see. Like how close are they, what part of it do they get to?" In the same view, Andrew also used the color demarcations to make inferences about which students to group together and which skills may be the focus of each group. He thought out loud, asking,

> Who's transitioning from one skill to the other that can be potentially one group? Um, I have a… you know group that understands or is ready for number lines. I have that. So it kind of goes back to that, how am I grouping my students? How? What kind of intervention strategies am I looking at?

While teachers like Andrew and Geneva focused on how the color code and sorting features of the prototype could be helpful in constructing homogenous small groups of students with similar levels of understanding, Natalie described how the prototype might facilitate purposeful pairings of students with mixed abilities. When asked how she might interpret the data presented in the class level view, without students ordered by score, Natalie walked through a potential process for forming small groups:

> I would probably use this data to create small groups even

> in a class. So, I would use it to maybe make some groups where these six students over here (draws a circle around the first six dots on the left) might be a good group to work together because you've got some students that are doing pretty well and they've definitely mastered the parts to the whole and they can help these ones to put some puzzle pieces together if you will.

In this way, Natalie saw the class-level view as a tool for creating groups in which certain students could serve as "good leaders" for other students working on different skills. This insight is important both because it diverges slightly from the commentary of teachers like Geneva and Adam, who anticipated using the prototype to group students by like-skills, and because using the prototype to create purposeful pairings fit into Natalie's current pedagogical priorities. Helping students "teach" each other was, she explained, her "super focus right now, that's what [she's] been working on the past year."

Comments like Natalie's, Andrew's, and Geneva's illustrate the variety of ways teachers envisioned using the prototype to plan and deliver small group instruction. As interviewers, we never mentioned the potential relevance of the prototype for instructional grouping, but nearly every teacher spoke about grouping as a way to use the information presented in the prototype in their practice (see Table 3). Because incorporating differentiation and intervention strategies tends to be a pressing instructional goal for individual teachers and schools, especially as they continue to recover from the interruptions of the COVID-19 pandemic, it was exciting to hear teachers talk through ways the prototype would support them in tailoring small group instruction to the needs and strengths of individual students.

Use of prototype to communicate with students and parents

The idea that teachers could use the prototype to communicate with students and parents was a theme that emerged during four of the interviews (positive in all cases). Teachers imagined scenarios in which a version of the prototype containing student-friendly language could motivate students and support a student-centered goal setting process. For example, Geneva offered the following suggestion,

> I'd like it (the prototype) in the type of language, like, students use. Because for me this was really helpful (reads from bulleted list of Level 4 description), 'when students use ratios as multipliers,' but something like, friendly language for them so we can even evaluate with the students: so maybe this is where you struggled and this is where we need to work on, work at. By the end of this lesson, you should be able to…

Geneva was thus picturing how the existing descriptions of each LP level might be transformed into a student-facing resource that could help students themselves understand their own strengths and areas for growth. Catherine echoed this idea by describing how she currently conferences with students about their growth goals for upcoming administrations of the i-Ready Diagnostic and wondered aloud about how the Student View of the prototype–an example of which is found in Figure 8–might fit into that process.

Figure 8. *Individual Student View with Grade Alignment*

She said, “So this is very interesting, if we have that conversation and then we could see a graph like this (Student View), we could call the student in and say see how much growth you’ve made!” Catherine and Geneva thus both started to propose ways information from the prototype might be leveraged to engage students in their own learning by setting goals and celebrating growth.

While Catherine and Geneva focused on the prototype’s utility as a tool for individual conversations with students, Natalie saw opportunities to incorporate the prototype into whole-class discussions about progress and help her students get "a little more motivated and excited." She brainstormed,

Honestly if there was a way that I could get rid of student initials entirely so that you

> couldn't tell who's who at all, I might use this (Class View) to show them where they are at as a whole, as a class. And say, okay, we as a class are great at these types of problems (clicks on and circles the example items for Part Whole and Fair Shares), but what can we do to start making sense of harder questions like this (circles example item for Multiply and Divide)? Um, just to kind of get them thinking about it a little more and get them a little more motivated and excited, maybe to get some growth because that's something that's just, something, you know, their growth mindset, you know, getting them to take responsibility and confidence in what they're doing.

Whether to paint a picture of whole-class progress or to facilitate one-on-one conversations with individual students, teachers believed the prototype held promise for motivating students and talking transparently with students about their learning. Because it was not a response we anticipated when we created the interview protocol, we were particularly interested to hear the reflections of teachers like Natalie, Catherine, and Geneva.

In addition to pointing to how the prototype could facilitate individual or whole class conversations with students, teachers shared thoughts about how the prototype might improve parent-teacher communication. For example, Danielle commented that the ability to quickly display a particular range of scores would be "especially helpful when I have to justify things for parents," such as whether a student would qualify to be tested for gifted and talented services. Catherine echoed this idea, explaining that having the detailed learning progressions document, which is accessible by clicking the question mark icon in the prototype, would make it "extremely easy to provide parents with information about what scores mean" in a parent-teacher conference. These types of comments indicated that teachers saw the potential for the prototype to function as more than merely a means of reporting test results to parents. Rather, teachers imagined that features of the prototype could contribute to substantive

conversations with parents about how their student is understanding a topic like fractions and how that understanding is reflected in the results of the i-Ready Diagnostic.

Use of prototype for teacher professional development

Five teachers described interest in the prototype's potential as a tool for teacher learning and development. Comments in this area focused on the following features: example items, LP level descriptions found when clicking the buttons on the right hand side of the prototype, and the detailed learning progression document (visible when clicking on the ? button next to the learning progression; see Figure 9 and Appendix B).

Figure 9. *Classroom View with "?" Option for Detailed LP View*

Several participants suggested that newer teachers or teachers who were less familiar with math and vocabulary standards could use these elements of the prototype, especially the detailed LP document, to deepen their content knowledge. Rachel, for example, was optimistic that teachers would be able to use the document to "drill down . . . into the standards." Tapping into her role as an instructional coach who works with teachers at a variety of experience levels, Courtney explained,

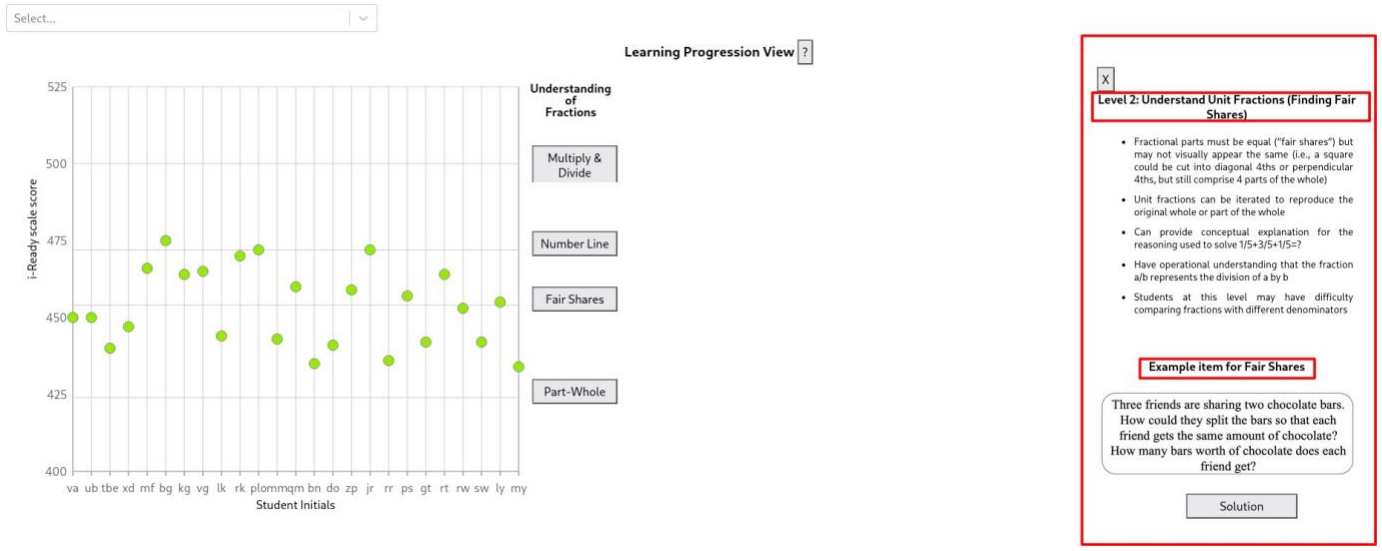
> A lot of the elementary teachers in our state, the way we do certification they're not really content specialists, so when we switched to Common Core math it was a big change for people. We had a lot of veteran teachers who were used to just teaching algorithms and rules. So, fractions has been kind of hard for a lot of our teachers. In saying that, *I* understand a lot of the progressions and the differences between them but I think it could be a challenge for someone where math is not their strength.

This statement from Courtney helped contextualize a challenge confronting teachers and instructional leaders and how the prototype might fit into that challenge: teachers who feel that "math is not their strength" may have a particularly difficult time making sense of academic standards and other materials that outline how students progress in their math understanding, and an LP like the one in the prototype could be supportive for such teachers. Geneva elaborated further, suggesting, as Courtney did, that "the standards vocabulary can be very specific," and the more detailed LP document, with its bulleted lists of common errors, "is easier to understand, especially if you're newer to the materials." Together, Courtney and Geneva's reactions suggested to us that the information shared in the prototype–level descriptions, example items, and the more detailed LP document–holds some promise for helping teachers learn more about topics in math instruction and math standards.

Catherine commented that some teachers, as they continue building their math content knowledge, would benefit from concise explanations of precisely what skills students at a particular level need to acquire and what kinds of problems they ought to be able to solve. After examining an example item and level description (see Figure 10), she remarked that

> This is the type of question that I give when I am interviewing for a tutor. This is the question I give them because this requires conceptual understanding, not the memorization of what we do to divide fractions. This is the essence of how we want our teachers to teach . . . This is a perfect question.

Figure 10. *LP Level Description and Exemplar Item*



She elaborated on this point later in her interview, referencing the bulleted list in the level description and exemplar item (see box depicted on right hand side of Figure 10): “This is, I think, what teachers need to know. If that's what you want [students] to do, this is how you have to teach it . . .”

This statement suggests that Catherine saw congruence between the information contained in the prototype and her vision for how she, as an instructional leader, hopes teachers on her team will deliver math instruction. Referencing that congruence, she reflected that the prototype "is important because it just verifies that what I am telling them [teachers] is the truth." It was encouraging to hear that instructional leaders like Catherine viewed the prototype as aligned to the kind of exemplary teaching practice they work toward in their schools.

Discussion and Future Directions

The research questions for this study were:

1) To what extent can an LP approach to growth, as represented by the prototype, facilitate meaningful interpretations about differences in student scale scores on the i-Ready Diagnostic?
2) To what extent are teachers able to understand and interpret the LP by using the prototype?
3) How do teachers envision using information provided by the prototype?

The interviews that formed the basis for our analysis provided insights relative to all three questions. The deductive portion of our coding provides encouraging evidence that a broader audience of teachers may easily and appropriately interpret student growth in the way intended by the prototype, which sought to highlight the relationship between the i-Ready scale and the fractions LP. The inductive coding process revealed several answers to the more open-ended third research question: teachers considered using information within the prototype for student grouping, professional learning, communication with students/parents, and more. Overall, the seven interviews provided an

abundance of evidence that we are on the right track and that CRG-based reporting would be useful to educators in the field.

We found that teachers were generally enthusiastic about both subjects we asked about (deductive codes) and themes that emerged independently of our questions (inductive codes). Tables 2 and 3 show that positive evidence was more common than negative evidence on every theme. The prototype was particularly successful in supporting teachers' connection of scale scores to the LP, as well as interpretations related to growth. We take the general enthusiasm we found as evidence that we should pursue further refinements to the CRG prototype in order to support users in ways suggested by our interview participants (we did not, for instance, find any indication that the basic idea of CRG needs to be reconceptualized, or that it would be overwhelming for potential users).Thanks to our interviews, we have arrived at several next steps that build upon this initial, largely successful effort.

At one point during initial development, the CRG reporting prototype provided two views: (1) a class-level view across the entire Fractions learning progression, which is now the only class-level view in the live prototype; and (2) a so-called "zoomed-in" view targeting the specific portion of the scale where students in the class are currently located, which we opted to remove for these interviews. Based on feedback that focused on the connection between the prototype and information about "standards," as well as concerns about the relative difficulty of items at the Fair Shares and Number Line levels, we are now working toward a revised approach to "zooming in" that could address both these areas of feedback. The first significant change from the prior zoomed-in view is that this view would now zoom in on a specific level of the learning progression. This would involve restricting the y-axis according to limits defined by the distribution of difficulty of items tagged as corresponding to the given learning progression level–for example, the 20th and 80th percentiles of difficulty. Accompanying this zoomed-in

view, we would present users with several example items for the given level, rather than just one. This would address feedback that the Number Line item appeared easier than the Fair Shares item, as the use of multiple items should make it clear to users that (a) there is quite a bit of variability in item difficulty *within* each level, and that (b) one can ask questions at each level of the learning progression that use simpler or more complex item layouts and tasks.

Additionally, in the zoomed-in view, we plan to build on users' enthusiasm for our association of scale locations with qualitative description of learning progression levels by incorporating information from the Common Core State Standards in Math (CCSS-M). Because the view is zoomed into a smaller subset of the i-Ready scale, there is likely enough room on the screen to locate specific standards from the CCSS-M along the right side of the y-axis, analogous to how we locate levels of the learning progression as a whole in the current view. We expect that strengthening the connection of the CRG reporting view to the CCSS-M would further strengthen educators' feeling that CRG is something that makes their lives easier and helps them understand more about what their students know and can do.

Beyond this, there are at least two other areas that we will need to address in the next phase of prototype development. First, it was clear that the probabilistic relationship between students' scale scores and the learning progression level in terms of anticipated performance was not clear until we explained it–several educators, for example, interpreted students' locations as corresponding to "mastery" of the learning progression level in which the student is located, when in fact the location corresponds to a 50% chance of getting an item at that difficulty correct. Educators' overall feedback on this was that the most important thing to do is to make that relationship clear to end users, and this is an important area for improvement of the prototype.

Second, we should likely revisit the Part-Whole and Fair Shares LP levels, independent of our efforts to reintroduce a view that zooms into each level. We received feedback that the move from Part-Whole to Fair Shares may represent too large of a jump for students. We should revisit the literature and empirical evidence on which the Fractions progression was originally based to see if there is a sensible and theory-backed way to split the Fair Shares level, in line with one educator's feedback, or perhaps to introduce a new level between Fair Shares and Part-Whole in line with another educator's suggestion.

To summarize, this small study represents a significant step forward for the CRG project. The response to our work among educators we interviewed was generally very positive, and there appears to be a real appetite for the type of information we propose reporting, alongside current reporting practices.

References

Clements, D. H. & Sarama, J. (2004) Learning trajectories in mathematics education, *Mathematical Thinking and Learning*, 6:2, 81-89.

Kennedy, B.L. and Thornberg, R. (2018) Deduction, Induction, and Abduction. In U. Flick (Ed.), *The sage handbook of qualitative data collection*. SAGE Publications Ltd, https://doi.org/10.4135/978152641607

Snyder, C. M., & Delgado, H. P. (2019). Unlocking the Potential of Data-Driven Coaching: Child Assessment Evidence as a Guide for Informing Instructional Practices. *YC Young Children*, *74*(3), 44–53. https://www.jstor.org/stable/26789001

Wellberg, S., Briggs, D., and Student, S. (2022). Big Ideas in the Understanding of Fractions: A Learning Progression. Center for Assessment, Design, Research and Evaluation.

Appendix A

Teacher Interview Protocol

Research Questions:

(1) How can a learning progression (LP) approach to growth, as represented by the fractions LP prototype, **facilitate meaningful interpretations of scale scores** on the i-Ready diagnostic?
(2) How are teachers **able to understand and interpret the learning progression** by using the prototype?
(3) How do teachers **envision using information** provided by the prototype?

Purpose: To solicit feedback from teachers about the…

- potential **usefulness** of embedding LP information into the i-Ready diagnostic report.
- **substantive interpretations** supported by the LP prototype
- **prototype's usability**
    - Note: we expect CA to adjust/enhance the visual aspects of the prototype. However, our goal is to build out the functionality and information that is conveyed in the interactive tool.

Plan:

- Individual 1 hour long interviews with 8-10 teachers in grades 2 – 5 who teach math

Agenda:

Pre-Work: Screener Survey

- Introduction
- Teacher think-aloud for LP feedback
- Wrap Up

Intro Script:

Thank you for agreeing to join us! We're looking forward to getting your feedback about how we can make our work more useful for you, as a classroom teacher. We wanted to start with a brief overview of who we are, what we've been working on, and goals for what we are doing today. Please feel free to pause us at any point if you have questions or would like clarification. We'd also like to pause here to ask if it's okay for us to record this Zoom. We'll use the recording, along with our notes, to revisit ideas that come up in our conversation today. Is it okay for us to start recording?

Who We Are

*First, we're going to introduce ourselves (*individually introduce who's on the call). *We are currently working in partnership with Curriculum Associates to help make their diagnostic assessment more useful and actionable for teachers.*

Survey Follow Up

Given that goal, we are interested in knowing: How, if at all, do you currently use the i-Ready Diagnostic reporting system? (Follow up: how frequently do you reference reports? How useful do you feel the reports are?)

Goals for Today

Today we're going to show you a prototype of a reporting experience that our team is designing. The hope is that this experience makes assessment data more useful and actionable. We want to understand how users feel about the new experience, which is why we're getting feedback from educators like you. The reporting experience is structured around Learning progressions. Learning progressions lay out the path students typically take in learning big picture concepts in math or reading. The prototype we've created connects a learning progression on an example topic, fractions, to students' performance on i-Ready assessment questions. We're going to ask you to interact with the prototype, and we'd love for you to poke around so we can hear your reactions and learn how to make it better.

Teacher Feedback Session:

Now we're going to have you interact with the prototype. Feel free to pause and let us know if you need to take a break for any reason, or have questions at any point.

I'm going to open the prototype, then share my screen and give you mouse control through Zoom. (Use the mouse icon in the top bar of Zoom to give mouse control to participant). **Let's do a quick test to make sure you have control of the screen.**

1) Start with LP-level view, no example item displayed, and no dropdown options selected.

Let's start with this. Before you click around and start to explore, which we'll invite you to do in a second, could you start by just narrating what you're seeing and what you think it might mean or represent?

Follow Up Prompts: How familiar are you with the i-Ready scale score? How do you interpret that here? Do you see anything you might click on that might help you understand what the far right axis means?

Go ahead and click on anything that you see and want to explore. As you do, please narrate out loud what you are thinking. For example, you might say, "I'd like to be able to…" or "I wonder what this means" as you click on different features of the prototype.

a. Prompts for think aloud narration: **What do you see? What does it make you think about? What do you want to do next?**

(Provide 6-7 minutes for participant to click around on different tools/ functions)

2) How might you interpret the data presented here about this class of students?

Score Uncertainty Follow Up Prompt (if teacher asks about score uncertainty range): Every time a student takes this test, they take one set of questions. The score uncertainty range shows you how the student might have scored if they'd gotten a different set of questions. For any subset of questions a kid gets, they could have gotten, a different score, and that's what we're trying to capture with this score uncertainty range.

3) How might you find more information about what a particular LP level indicates about student understanding?

a. Do the differences between the levels of the LP make sense to you?
Follow Up Prompt: Feel free to take your time and review the levels from the gray buttons.

b. Can you reflect on how useful (or not) the written description and exemplar item are in helping you to understand what the LP Level means?

c. Now let's look at a sample student, the student with the initials XD. How do you think you might find more information on XD's performance?

d. From the data you see here, how would you describe XD's understanding of fractions?

e. How would you describe XD's growth over time?

4) **In your instructional planning or practice, how might you use what you've learned from this class-level view, if it all?**

5) (If teachers have not yet navigated to the dropdown, "Color Code Learning Progression" , ask…). **We're curious what you think about including color in the presentation of the learning progression. What are your reactions to the drop-down menu called "Color Code Learning Progression"?**

6) If you were going to tell a colleague about what we looked at together today, what would you say?

<u>Wrap up:</u>

Now that you've had a chance to see the entire report, we would love some of your general feedback. (Select 2-4 of the questions below, depending on time)

7) Was there anything that you particularly liked or didn't like about the tool?

8) Is there anything you would like to see that is missing?

9) What questions do you still have about what we looked at together today?

10) What advice do you have for us as we continue to iterate on this work?

11) Is there any more information you feel like you need to understand the fractions learning progression better?

12) Knowing what we've told you about our plan for this project and what you've seen today, is there anything else you'd like to share with us? Is there anything you think we should have asked you about that we didn't?

That wraps up our initial user test. should be receiving your honorarium in the form of an Amazon e-gift card from Curriculum Associates' User Experience Operations Manager. We'll put her email address in the chat, so you have it to reach out if necessary. We also wanted to ask: Would you be interested in hearing from us with updates about the progress of this project? Thanks so much for participating today! We really value your feedback.

Appendix B

Detailed LP Document

Fraction conceptualizations learning progression (LP)

LP Level	Student Characteristics	Item Responses
4. Multiply & Divide	**Understands that:** • Multiplying a value by a fraction $\frac{a}{b}$ results in a value that is a-$b$ths of the original value • Understands the difference between multiplying and dividing fractions	**Is able to:** • Use multiplication to find a portion of a value • Determine that multiplying a value by a fraction with magnitude less than 1 will result in a value with smaller magnitude and multiplying by an improper fraction will result in a value with larger magnitude, and vice versa for division, without performing the calculations • Divide a value by a fraction **Common Errors:** • May not understand the conceptualization of a fraction as a ratio • May have diffculty understand practical applications of fractions in solving real-world problems
3. Represent on Line	**Understands that:** • Fractions represent unique numerical values (real numbers) • Two fractions are equivalent if they represent the same numerical value • Fractional values can be converted to decimals or percentages while maintaining their numerical value • Improper fractions may be rewritten as mixed numbers and vice versa • Fractions with different denominators may be readily compared, added, or subtracted once they are put into the same units **May not yet understand that:** • Fractions may be written as ratios and may represent part-part relationships or rates	**Is able to:** • Create and identify equivalent fractions, including converting between improper fractions and mixed numbers • Order fractions and mixed numbers with different numerators and different denominators • Add and subtract fractions and mixed numbers with different denominators **Common Errors:** • May not understand the conceptualization of a fraction as a ratio • May have diffculty understand practical applications of fractions in solving real-world problems
2. Understanding Unit Fractions (Finding Fair Shares)	**Understands that:** • Fractional parts must be equal ("fair shares") but may not appear the same	**Is able to:** • "Share" a whole between a specified number of groups • Identify unit fractions