ZDM Mathematics Education (2015) 47:919–932
DOI 10.1007/s11858-015-0699-y

ORIGINAL ARTICLE

A design research study of a curriculum and diagnostic assessment system for a learning trajectory on equipartitioning

Jere Confrey[1] · **Alan Maloney**[2]

Accepted: 16 May 2015 / Published online: 16 September 2015

Abstract Design research studies provide significant opportunities to study new innovations and approaches and how they affect the forms of learning in complex classroom ecologies. This paper reports on a two-week long design research study with twelve 2nd through 4th graders using curricular materials and a tablet-based diagnostic assessment system, both designed around a learning trajectory on equipartitioning. A learning trajectory is a conceptual model of how students move from naïve to more sophisticated understandings as they engage with a carefully sequenced set of tasks. The equipartitioning construct describes how students come to understand the ideas involved in sharing fairly an evenly divisible collection, a single shape or multiple shapes. The paper is organized around the three phases of design research: planning, conduct, and retrospective analysis Cobb et al. (Educ Res 32(1):9–13, 2003). It illustrates how the conjectures of the study are subjected to testing and revision, based on the students' and teachers' behaviors during the study, and how interpretations and theories evolve during the different phases.

Keywords Design research study · Learning trajectory · Diagnostic assessment · Equipartitioning · Prospective and retrospective analysis

✉ Alan Maloney
maloneyap01@gmail.com

[1] North Carolina State University, Raleigh, NC, USA

[2] Cary, NC, USA

1 Introduction

In early grades, young children concentrate on learning counting, place value, and addition and subtraction to the neglect of building an adequate foundation for division, multiplication, ratio and fraction. Confrey (1994) proposed an alternative emphasis in the "splitting conjecture": splitting (fair sharing) constitutes an independent but equally important foundation for building students' mathematics understanding. To address this other foundation, Confrey et al. (2009) proposed the concept of equipartitioning, "cognitive behaviors that have the goal of producing equal-sized *groups* (from collections) or *pieces* (from continuous wholes) as 'fair shares' for each of a set of individuals," and a related learning trajectory that outlined the expected behaviors of young children as they progress from naïve to more sophisticated understandings (Confrey et al. 2014b). To facilitate a more widespread implementation of equipartitioning in schools, the research team built a learning trajectory-based set of curricular materials and a diagnostic assessment system delivered on tablets.

The paper reports on a two-week design research study using the curricular materials and diagnostic assessment system with young children (ages 7–9 years) learning the concepts of equipartitioning of both collections and wholes. The curriculum materials were built around the equipartitioning learning trajectory (EPLT). The interactive diagnostic assessment system or IDAS (Confrey et al. 2011a) was designed to record and evaluate student strategies in solving equipartitioning tasks in real time, and provide rapid, periodic, and precise feedback to students and teachers concerning students' progress in the EPLT. The study examined the combined effects of the curriculum, instruction, and diagnostic assessment system on student learning. Based on the explicit theory of cognition (the EPLT), the team had expectations of what was likely to happen, but new materials and new technology made it necessary to test those expectations against real practice.

The central research question was:

What do students learn about equipartitioning from using the curricular materials together with the interactive diagnostic system?

Subsidiary research questions were:

- *To what extent can children, ages 7–9, learn to share collections and single wholes accurately and to name, justify and reassemble those results as predicted by the lower levels of the equipartitioning learning trajectory? Which concepts are easy or difficult for the students to learn?*
- *What effects on classroom learning can be observed or conjectured, from use of the interactive diagnostic assessment tool? How do the results from the diagnostic tool support students' discourse and what does it reveal about their understanding of the selected concept of equipartitioning?*

2 Methodology

A design research methodology was adopted as a powerful research tool for introducing newly emerging ideas, tools, and technologies into schooling, and studying their effects on student–teacher interactions and learning distinct from experimental studies in psychology (Brown 1992; Design-Based Research Collective 2003; Collins 1992; Confrey 2006). It permitted the research team to undertake "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). Design research also evolved in contrast to curriculum research aimed at establishing comparative effects of new curricula. It was designed to gain insight into student learning in the complex environment of student and teacher interactions, and to document the importance of local and domain-specific instructional theories (Streefland 1984, Gravemeijer 1994, van den Akker et al. 2006, Artigue 2008, Plomp & Nieveen 2013). By specifying the research questions more finely in the form of conjectures, design research studies permit one to pose tasks to create observations of related behaviors, and to refine those tasks based on student responses.

Design research has both pragmatic and theoretical orientations. Their goals are to create practical design solutions to important challenges, subject those solutions to careful examination and revision, *and* make significant contributions to learning theory. This study was highly interventionist (as are many design research studies), designed to develop understanding about how our new curriculum and prototype IDAS could be incorporated into instruction, and its effects on classroom interactions. The study involved a complex learning ecology whose elements included "the tasks or problems that students are asked to solve, the kinds of discourse that are encouraged, the norms of participation that are established, the tools and related material means provided, and the practical means by which classroom teachers can orchestrate relations among these elements" (Cobb et al. 2003, p. 9). Taken together, the EPLT, the curriculum materials developed for it, the IDAS, and classroom practices responding to repeated and just-in-time data use, represent major components of a new learning ecology that can be studied as it is manipulated.

The current report comprises three main sections, illustrating three phases of the design research study. *Preparation* focuses on the background to the study. *Prospective and reflective aspects* deal primarily with how two conjectures on student learning were examined in relation to the curriculum and the IDAS. *Retrospective analysis* uses data from the IDAS to reflect on the research questions.

3 Preparation

3.1 Equipartitioning learning trajectory

Learning trajectory research focuses on delineating the patterns of behaviors exhibited by children as they learn key ideas, typically over multiple years. They have been articulated for early concepts in geometric shape (Clements et al. 2004), number sense and operations (Clements and Sarama 2009), measurement (Barrett and Battista 2014), data modeling and statistical reasoning (Lehrer et al. 2014), and fractions (van Galen et al. 2008) among other topics. They describe children's early ideas typically rooted in common experience, a knowledge target based in the discipline, an articulation of the common landmarks and obstacles, and ways to address them. For equipartitioning, the early common experiences are fairly sharing collections and single whole objects; the target understandings are the various meanings associated with the idea that one quantity, a, compared to another quantity, b, is equivalent to a/b compared to 1. Its significance lies in relating division, multiplication, ratio, and fraction to the operation of creating fair shares (equal-sized groups).

The EPLT (Confrey et al. 2009), this study's underlying theory of cognition, was developed through synthesis of mathematics-education and cognitive psychology research, numerous additional cross-sectional clinical interviews, and field testing of 120 paper-and-pencil field test items with approximately 5000 K-8 students (Confrey et al.

ePacket #s ***EP LT Proficiency Levels***	**P1**	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**
16: Generalization: a among $b = a/b$							
15: Distributive property, multiple wholes							X
14: Co-splitting							X
13: Composition of splits, multiple wholes						X	X
12: Equipartition multiple wholes						X	X
11. Assert Continuity Principles					X		
10. Property of Equality of Equipartitioning					X		
9. Redistribution of shares (quantitative)			X				
8. Factor-based changes (quantitative			X	X			X
7. Composition of Splits; factor pairs				X			
6. Qualitative Compensation			X	X			X
5. Reassemble: n times as much as…		X	X	X			X
4. Name a share with respect to referent unit		X	X	X			X
3. Justify the results of equipartitioning	X	X	X	X	X		X
2. Equipartition single wholes		X					
1. Equipartition collections	X						

Fig. 1 Equipartitioning learning trajectory proficiency levels (*1–16*) mapped (X's) to curricular and diagnostic packets *P1–P7* (Confrey et al. 2014a)

2014b). It comprises 16 proficiency levels (Fig. 1). For each level, task classes (an ordered set of numeric parameters, not shown here) facilitate systematic variation in the types and difficulties of problems (e.g. in sharing a single whole, most children find equipartitioning for 2, 4, and 8 persons easier than for 3, 5, or 7).

3.2 Curricular materials

Paper-and-pencil curriculum packets were developed (Confrey et al. 2011b), distributed across seven packets as shown in Fig. 1.

The curriculum packets address the following topics:

P1) Equipartitioning collections (e.g. 12 candies shared among 3 children)
P2) Equipartitioning a single whole (e.g. one pizza shared among 4 children)
P3) Reallocation (adapting equipartitioning strategies to changed conditions, e.g. after 24 items are shared among 4 children, one child departs, so the collection is re-shared among 3)
P4) Composition of splits (anticipates multiplication and factors, e.g. a rectangular cake is shared among 6 by making a 3-split in one direction, and a 2-split in the orthogonal direction)
P5) The property of equality of equipartitioning (PEEQ; e.g. two congruent rectangular cakes are split into 2 equal parts, one horizontally, the other diagonally; student recognizes that while non-congruent, all the pieces are equal-sized)
P6) Co-splits of multiple wholes (e.g. 12 baseball players order 8 pizzas, but cannot sit at a single table. Students identify table configurations that maintain ratio of players to whole pizzas).
P7) Equipartitioning Multiple Wholes (e.g. 5 pizzas shared fairly by 3 children).

Proficiencies of justifying, relational naming, and reassembly of shares into the original whole (Fig. 1) were incorporated into multiple packets.

During the study reported in this paper, students worked with packets 1 and 2, and therefore with EPLT levels 1 through 5.

3.3 Interactive diagnostic assessment system (IDAS): *LPPSync*

Strategic placement of formative assessments within instruction often improves student outcomes (Heritage 2010; Popham 2008; Black & Wiliam 1998), but typically fails to result in systematic record of student progress over time. A precise diagnostic assessment, delivered in real time, could serve the purpose of informing instruction while providing specific reports to help students themselves know where to concentrate, and class-level reports to support teachers understanding the distribution of their students' proficiencies on the learning trajectory. The team developed an interactive diagnostic assessment system (IDAS), named *LPPSync*[1], to run in browsers on iPads (Confrey et al. 2011a; Confrey & Maloney 2012).

The IDAS can be used in two activity modes: diagnostic (assessment) and practice. In diagnostic mode, students

---

[1] Learning progress profiles synchronized for networked wireless devices.

receive six items to solve. The software scores their performance in real time and generates feedback to the student and the teacher. In practice mode, students can choose the item difficulty, and receive feedback as they solve each item. Practice mode includes "chat," a social communication feature to support collaborative work (students can text, share screen images, and observe each other's screen actions in real time).

LPPSync generates items of varying difficulty, drawing from a parameterized sample space based on the task classes. In a diagnostic assessment, items are generated at easy, medium, and hard levels, two per level. The LPPSync database records students' responses, including their strategies and justifications, and generates reports. For each curriculum packet, LPPSync deploys "virtual manipulatives," on-screen objects and tools that students use to solve problems. Actions on the virtual manipulatives are recorded as data by the software. In a diagnostic assessment, students work individually to respond to all six randomly-generated items. In practice mode, teachers and students can form student groups, collaborate and communicate about solutions to items, specify the parameters of tasks, and try multiple solutions to a task.

Diagnostic assessments were used formatively, i.e. pre-assessments were administered prior to instruction on the packet, then again (as a post-assessment) after the instruction on the topic was completed; they could also be used at other times if an instructor determined they would be useful. Practice mode items, activities, and tools were incorporated into instruction as planned by instructors. We emphasize that we were not investigating the "value added" of the IDAS as might be done in a comparative or controlled study. Instead, we sought to understand how the IDAS could interact, as a tool for formative assessment and instructional support, with the learning trajectory and curriculum materials, in the overall learning ecology, as a means of theory generation about the role of a diagnostic assessment software tool within instruction based on a specified cognitive framework. For example, what should be the timing, frequency, and substance of feedback from a diagnostic assessment, to strengthen student learning?

3.4 Detail on packets 1 and 2 and related diagnostics

Packet 1 (curriculum and LPPSync) combined LT levels 1, 3, 4, and 5: fair sharing of collections of objects, justifying the results, naming the shares in relation to the collection, and describing the size of the collection in relation to the size of the share (reassembly), to strengthen understanding of the inverse of equipartitioning. Students are first expected to develop strategies to fairly share collections. The specific task is to divide up a set of children into teams to play games requiring a variety of numbers of participants. The correctness of fair sharing a collection can be justified either solely using one-to-one correspondence between the members of each share or by counting to see that the shares are equal: one can conceive the result of sharing a collection as either producing a certain number of objects per share or as producing a certain number of fair shares relative to the whole. In clinical interviews, children commonly reflect this distinction as they name the results of fair sharing (e.g. for 10 objects shared by two children) in two ways, as a count ("we each got 5") or as part of the whole collection ("each got half"). The EPLT identifies these two ways of naming a share as: 1) a *count* of the "number of objects" per person, and 2) *relationally*, as one-*p*th of the whole collection (*p* equal to the number of people sharing) as a description of "relative size". In the case of sharing a whole, only one form of naming applies—the relational naming of a share as a unit fraction (1/*n*th of 1). Relational naming represents early foundations among ratio and fractions, and notions of equivalence.

Packet 1 tasks also require students to be able to describe the size of the whole collection relative to the size of share, in terms of how many times as large as, or as many as, it is than the single share, emphasizing the importance of reversibility (Piaget 1970) in establishing the robustness of an operation such as equipartitioning. Furthermore, because the act of fair sharing evolves into division, the act of *reassembly* (as named in the EPLT) will support early forms of multiplication.

Diagnostic items in packet 1 were designed for students to conduct equipartitioning actions on virtual manipulatives; if successful, the student was then asked to name the shares and collections in relation to each other and to the number of sharers. The diagnostic naming questions were as follows (e.g. 12 coins shared fairly among 3 people):

(a) "*How many* coins does each person get when 12 coins are shared fairly among 3 people?" (naming by count)
(b) "*How much of* the whole collection does each person get?" (relational naming)
(c) "*How many times as large as one share is the whole collection?*" (reassembly)

Packet 2 focused on the strategies students developed when sharing single wholes. The tasks involved equipartitioning different shapes using different numbers of colors. LPPSync items involve simple 2-D shapes (single wholes), and tools to mark the shapes, adjust marks' positions, cut the shape into pieces, to translate, reflect, and rotate pieces to test for size and shape equivalence, and color pieces to distinguish shares of the original whole. The diagnostic system recorded the students' methods and the area of the parts (to evaluate equivalence of parts) (Fig. 2).

The design research study would also provide feedback on the students' user experience in answering these questions.

Fig. 2 LPPSync packet 2 task, circle shared among 4

The design had to accommodate variation in prior knowledge and reading ability, as well as use of keyboard entry. The LPPSync design required students to choose a response-type (whole number, fraction, text) from a pull-down menu, and then complete a corresponding answer field ("whole number" had a single-value entry field, "fraction" a two-value fraction entry box, "text" had entry from a keyboard).

3.5 Participants

Student participants (two second-graders, two third graders and eight fourth graders[2]) were from a local urban community center (whose staff noted that all the students needed substantial learning assistance). Research project staff noted early during the study that they demonstrated very weak study skills, highly variable social skills, and little experience working effectively in groups. They were reluctant to share their answers and exhibited fear of giving "wrong" answers. Considerable instructional time was devoted to strengthening the students' identities as learners, developing shared mathematical norms of listening and explaining (Yackel & Cobb 1996), trust of the research staff, and dispositions to persist and to work cooperatively. With support and positive feedback from the instructors they slowly began to understand the classroom norms. It took time before they were willing to volunteer to come to the board to talk about work samples.

3.6 Classroom schedule and organization

The teaching experiment comprised ten instructional sessions of 1.5–2 h over two consecutive weeks during school summer vacation. Two research team members served as instructors. The daily sessions included whole class, small group, and individual activities, and typically began with a whole group activity, review of the previous day's topic, or a diagnostic assessment. Diagnostic assessments were administered before and after the class worked on a packet. Children received their percent correct score immediately on completion of the assessment. Individual or group work was typically followed by whole-class discussions of student ideas. Research team members conducted clinical interviews with individual students on four occasions during the study. By the end of the 2-week study, the students had completed packets 1 and 2, and began packet 3.

3.7 Data collection

Data were collected continually throughout the study:

- video and/or audio-recorded class sessions and small-group interactions;
- written work artifacts collected in student notebooks;
- on-screen series of actions on objects, quantitative and verbal responses, and chat records were recorded in the LPPSync database;
- LPPSync-generated reports: mathematical challenge of each item, student item scores and percentage of diagnostic items answered correctly, breakouts of quantitative and qualitative responses including indications of common errors and misconceptions signaled by student responses (including use of virtual manipulatives), and group summaries.

4 Conducting the design research study: prospective and reflective aspects

Design research studies include conjectures (Confrey & Lachance, 2000) that link the research question(s) to the students' activities in the classroom setting. The conjectures facilitate observation of student behaviors and utterances around tasks and interactions, and help to identify, explain, or raise questions about variations. Unlike hypotheses, conjectures are not expected to be wholly confirmed or disconfirmed; rather they evolve over the course of the study—examined, refined, evaluated, modified, strengthened, or at times, rejected.

Prospective and reflective activity occurs daily—rounds of planning, reflecting, and revising. Our team met prior to, and debriefed after, each session. Debriefing sessions began with a free-form "brain dump," focusing on team members' observations and pressing on conjectures. Issues were sorted into broad emergent categories: "general topics," "curriculum," "software," "system reporting," and "interview topics." Each debriefing concluded with a revised plan and schedule for the next day's instruction.

---

[2] For family reasons, a 3rd- and two 4th-grade boys left the program after week 1; their work during week 1 is reflected in the study data.

We report here on two conjectures that represented key opportunities to test local instructional (Gravemeijer and Cobb 2006), or "humble" (Cobb et al. 2003), theories. The first conjecture focused on relational naming of fair shares and collections, the other on the criteria for successful equipartitioning. The design research study placed them in "harm's way," i.e. subjected them to iterative design, requiring them to inform "domain-specific learning processes" that are also "accountable to the activity of design" (p. 10).

The main research question concerned the overall impact of the materials and IDAS on learning equipartitioning. The sub-questions were to examine specifically how the students proceeded through the lower levels of the trajectory and the effects of the IDAS. The conjectures addressed specific instructional episodes that raised questions about the first two levels of the LT (equipartitioning collections and single wholes). At these levels, students are asked to demonstrate that they can create equal shares. The next three levels (3, 4, and 5) focus on how students name the results of those actions, justify their solutions, and whether they can describe what is obtained when those shares are reassembled into the original whole or collection. During the first day, it became clear that in the case of collections, naming the results yielded a variety of solutions. Hence, Conjecture 1 emerged.

Conjecture 1 *Children will readily adopt multiple ways of naming fair shares of collections, including counting and relational naming (naming one share in relation to the whole collection or single whole). For example, for* 12 *objects shared among* 3 *children, a share is* 4 *objects (per child). Relationally, each child receives* 1/3 *of the collection. Children will also be able to naming the original collection multiplicatively in relation to a single fair share using "times as many," or "times as much," e.g., the whole collection is* 3 *times as much as one share. The relationship between naming the size of the share and relating its size to the original collection will reinforce each other.*

When sharing collections, students relied on dealing (giving one or more items to each person, in sequence) as their main strategy, and were generally successful in correctly sharing a collection. Students would almost always deal to the correct number of people, deal all of the items systematically, and obtain the equal shares. If a round of dealing did not end on the last place correctly (atypical), students realized the need to begin again or count each pile to correct their errors.

As students moved to share a single whole, they often struggled to satisfy three criteria for equipartitioning (sharing for the correct number of people, making the share equivalent, and exhausting the whole), and Conjecture 2 emerged.

Conjecture 2 *Children's abilities to perform and reason about equipartitioning wholes become stronger and more flexible as they learn to explicitly apply the three equipartitioning criteria (correct number of shares, fair (equal size) shares, and exhausting the whole).*

Specifying conjectures as elements of a local instructional theory is a critical element of design research studies, requiring the researcher to track student understanding at a deeper level than might be suggested by the main research question. Conjectures set one's observational lens, heighten one's sensitivity to particular activities and instructional exchanges, and obligate the researcher to pose tasks and questions that push its limits. We report on how each conjecture played out across the days of the study.

4.1 The evolution of Conjecture 1 (naming fair shares)

During the first 2 days of instruction, the students readily learned to complete equipartitioning tasks with collections using manipulatives and LPPSync. After dealing the coins, they checked for the completion of the round and the equivalence of the shares. However, when it came to naming their results, it became apparent that the conjecture about the ease with which they name shares both as counts and relationally was in jeopardy. They regarded the questions "how many coins did each person get?" and "how much of the whole collection did each person get?" as identical; during class discussion and the first diagnostic assessment, most students answered both questions with a simple count of the items in a fair share (and found it challenging to respond to the questions in the software interface).

The research team developed an activity to try highlight the distinctions and relationships among the three questions ("how many objects?", "how much of the whole collection?", and "the whole collection is how many times as large as one share?") during day 3. We conjectured that if the n shares were each grouped into a container of some sort, so that the containers could be counted, the students might see how to name one share as 1 of n identical groups. To test this conjecture, we designated ten students as the items in a collection, and two hula hoops (large plastic rings) as the containers. A student "ringmaster" was assigned to sort his or her classmates into two "rings". In response to the question "how much of the whole collection is in each ring", each child said "5 out of 10." The team had expected a response of "half" from many students, based on previous clinical interview research. No amount of questioning or different combination of students and rings elicited the students to connect the "5 out of 10" to "half" or "one half." Using the rings to demark each share (group) and the number of shares (groups) seemed to help them to describe it more relationally using "out of" language, it did not lead to a description of one group in relation to two equal groups or the use of "one half." The team tentatively named this "5 out of 10" language a "proto-relational" naming strategy, because values of 5 and 10 could be expressed as a comparison of two counts of the share and the collection but also could indicate the beginnings of describing relative size or ratio.

Students also struggled with describing the whole collection as "times as many as" a share. They readily recognized that a collection shared among two was "double" the size of the share, but it was not clear that, to them, "double" meant "two times as many (or much)," for they did not extend "double" to "three times…," "four time…," and so forth: for more than two persons sharing, they tended to describe the collection in terms of how many *more* objects were in the whole collection than in the fair share.

We modified our conjecture again during the same class session to see if we could leverage doubling. Our goal was to have students first view a single share as "one times as big," then two shares as "twice as big," until they reached "*n* times as big" as the original share. We gradually expanded the size of the collection in relation to the original group, in a recursive way (including the original set each time) instead of an iterative way (counting the groups). The students worked up from an original share size of, for example, 3 coins (1 group) as "one times as much as" a single share, and then 2 groups as "twice as big as" the original share, with teachers encouraging students to replace "double" with "twice as large as." Then another share was added (three times as big), then another (4 times as big), and so on. This approach seemed to help students distinguish multiplicative ("times as big") from additive ("more than") language.

Debriefing after day 3, the team discussed students' preferences for answering the "how much of the whole collection" question proto-relationally ("3 out of 12,") instead of relationally ("1 out of 4"). We developed an equipartitioning activity using paper coffee filters as containers to delineate each share. We predicted that through the use of the filters (as containers) and class discussion, the students would name a share as 3 out of 12 *or* 1 out of 4, and that, just as they would describe one of two equal shares as ½ of the collection, they would adopt the relational language of a share as "one-fourth of the collection." We also predicted that, if they adopted the description of one share as "1 out of 4," they would more readily perceive the collection as 4 times as large as the single share. We did not expect them to make the further claim that $3/12 = 1/4$ at this juncture (equivalence of two fractional parts is a later development in a distinct learning trajectory, and builds, in part, on equipartitioning).

After some time with the activity (day 4), and discussing the ideas of "one *p*th of" and "*p* times as large," the students did recognize that the number of people sharing, *p*, corresponds to the denominator of 1/*n*th, and furthermore observed that this number was the answer to the "how many times as many as a share is the whole" question. The research team was uncertain about the degree to which these observations implied conceptual understanding, rather than the recognition of a pattern without understanding why. We also debated whether reproducing the pattern of language use was a step towards understanding (a pseudoconcept in Vygotskian terms) or merely a brittle rote memorization, and assumed a "wait and see" position.

On days 4 and 5, students also worked through most of packet 2, which focused on equipartitioning of wholes (next section). Unfortunately due to problems with the software and in the design of the interface for entering fractions, the data on naming wholes and reassembly in the software was unusable. In a paper assessment of sharing wholes on day 7, there was only one error in naming and reassembly across all students on all four problems showing eventual mastery of the idea.

On days 5 to 7, the team conducted clinical interviews to examine the interaction between students' understanding of equipartitioning collections and single wholes, and to explore the effect of working with wholes on their relational naming of shared collections. Most of the interviewed students readily equipartitioned collections and single wholes correctly, named shares of collections proto-relationally (e.g. 3 out of 12), and named shares of wholes relationally (e.g. one fourth). None of the students, however, spontaneously named shares of collections relationally. Probing further, two interviewees were asked if they could relate sharing a collection of 12 among 4, to sharing a rectangle among 4; both placed single shares of coins onto single shares of the equipartitioned rectangle. When asked how much of whole or collection the share was, they responded "one-fourth." The team subsequently conjectured that a primary challenge for students to see a share of the collection as 1/*n*th was to recognize the change in the referent unit for the collection from viewing a single coin as a whole to the collection as a whole. Instruction based on this conjecture will be tested in future study iterations.

This episode served an example of combining whole class instruction with follow-up clinical interviews, as another dimension of the flexibility of a design research study to allow a research team to pursue a conjecture and seek a possible resolution.

4.2 The evolution of Conjecture 2 (applying the three criteria)

Fairly sharing a single whole is the second major case of the equipartitioning LT, after sharing a collection. It draws on geometrical reasoning in which symmetries and congruence are at play. From prior literature (e.g. Pothier & Sawada 1983) and our own studies (Confrey et al. 2010), we

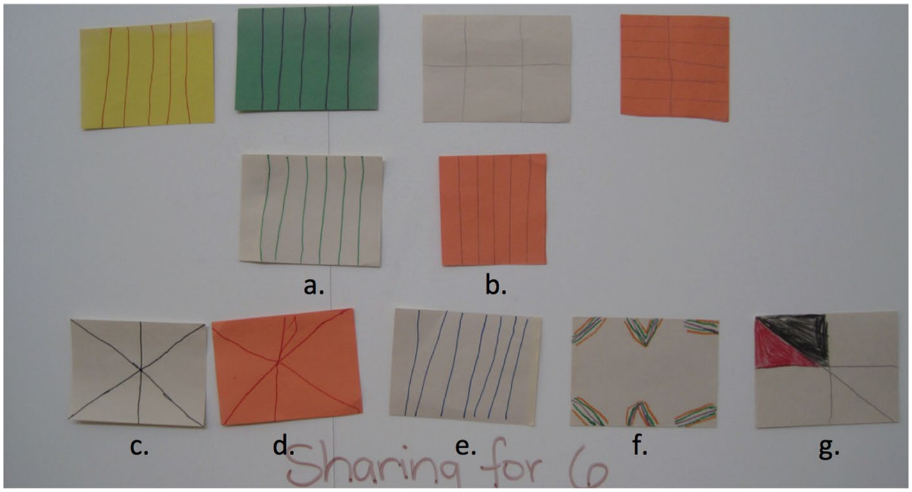

Fig. 3 Student work: sharing a rectangle fairly for 6

anticipated four common errors and misconceptions when students are asked to fairly share a single whole for p people.

1. p parallel cuts on a rectangle, resulting in $p + 1$ shares.
2. 3-Split strategies using a 2-split on the whole followed by a 2-split on one of the halves (three unequal pieces), or splitting vertically and horizontally (composing) to produce 4 fourths, and discarding one of them.
3. Vertical parallel cuts on a circle.
4. Composing splits—an a-split on side A of a rectangle, and a b-split on side B, but predicting the final number of pieces to be $a + b$ instead of $a \times b$).

In the packet 2 diagnostic pre-test (day 4), all but one student performed correct 2-splits on circles and rectangles; three successfully shared a shape for 4 people, and one shared a rectangle for 6 people. None successfully shared a shape for an odd number of people.

Day 5's activity aimed at identifying the criteria for successful equipartitioning. Figure 3 shows students' work on equipartitioning for 6. The top row contains correctly shared rectangles: examples of vertical cuts, a "composition of splits" combining vertical and horizontal splits, and a rectangle partitioned into 12 equal sized parts (2 pieces per share). The second row has rectangles on which 6 lines (7 parts) were drawn (*a*); one child "corrected" her work by cutting off the rightmost share to make six equal shares (*b*). The bottom row shows unequal sized pieces (*c*, *d*, *g*), too many (*e*) pieces, and a non-exhausted whole (*f*)[3]. Two students removed equal-sized pieces without using the entire whole; they explained that at home four children would never be allowed to eat an entire cake by themselves.

The class then discussed which features of the various equipartitioned rectangles and circles made them fairly shared (or not), to develop a set of criteria or rules. By the end of the discussion, most of the students could correctly invoke the three criteria to identify correctly and incorrectly shared shapes.

Progress on sharing for odd numbers of sharers was slow, especially for a circle, which many students claimed was impossible to share for an odd number of people. A few students began to view the cuts as radiating from the center and thought to rotate the radial cuts to make room for one more cut from a 4-split, to make a 5-split. A few children were able to draw shares for three, typically referring to a "peace symbol" or a Mercedes Benz logo when doing so.

However, a second administration of the diagnostic assessment did not show the gains we anticipated. So the team developed two activities to promote students' mastery of packet 2, and to strengthen their peer discourse. The diagnostic results were analyzed, and personalized worksheets for each student were created, with simple encouragements (e.g. "Great job sharing rectangles for 2 and 4 people. Practice sharing for 3 and 6 on a circle, sharing for 4 and 5 on a rectangle, and naming shares. Practice these problems to improve your score.") Each student completed four more problems. Nine students completed such personalized formative assignments, with only a single error, demonstrating either the power of personalized feedback from diagnostics or that the students indeed understood more than the LPPSync data showed[4].

---

[3] Some students struggled with the precision of the physical task; many said they had never tried to share make fair shares from single shapes.

[4] Only late in the study was it discovered that the internal software criteria for accepting equipartitioned areas as correct was inappropriately stringent: student responses had been scored incorrect for only minor variations in the sizes of partitions. This resulted in misscoring and denying students opportunities to answer most of single-whole naming and reassembly questions.

Fig. 4 Flags discussed by students in the examples

Second, a game was designed to bring LPPSync's chat feature into play to systematically encourage more mathematically oriented discussion, and to generate individualized data from collaborative activity[5]. National flags are excellent examples of partitioned rectangular spaces, varying in orientation, number, and "fairness" of partitions. The game challenged students to identify national flags that displayed particular criteria such as "equipartitioned vertically for three persons" or "three shares horizontally but not fair shares." Student partners, located in separate rooms, had to text to each other to decide jointly and justify solutions.

Students eagerly engaged in the activity. Younger students made only very simple exchanges, usually without justification. However, older students generated more detailed exchanges: the first example below is an excerpt from a conversation about the task "Identify a flag that is equipartitioned horizontally for 3 persons." (S: student. R: researcher. Flags referred to are shown in Fig. 4).

S1: I agree with Russia and Kenya, but not Surinam. Surinam green [is] smaller

R: What about Kenya?

S1: Yes bc [because] 3 square [rectangles}

S2 […] not a answer bc it has 5 parts and the white lines count

S1 I agree

The same students discussed "Identify a country whose flag is equipartitioned vertically for 2."

S2 No, there is no answer for that one

S1 None bc they all have more than 2

S1 Why

S2 The answer is Canada

S1 Canada has 3. Maybe if the white is twice the red. What do you think

S2 I agree and why [with your explanation].

These conversational exchanges show children using language of equipartitioning, paying attention to more than just the number of regions, and extending their notion of equipartitioning: S1 asserts that the Canadian flag satisfies the challenge if the central white stripe is equal in area to the area of the two outer red stripes combined, and uses a basic multiplicative comparison ("twice").

The students' sharing of wholes suggested that too often children in school are only asked to identify fractional parts of partially shaded shapes and are not provided enough opportunities to engage in the operation of equipartitioning. Only by actively equipartitioning do students begin to recognize interesting interactions between a shape's properties and the number of pieces into which it can be readily equipartitioned.

The design research study methodology was useful in examining Conjecture 2 in a different way than Conjecture 1. The classroom activity appeared to confirm Conjecture 2. Although the results of the diagnostic questioned that assumption, this provided an opportunity to leverage formative assessment practices as part of the learning and teaching ecology, to provide students with personalized opportunities to address their misunderstanding. This activity seemed to yield the desired outcomes. The design research study format also allowed researchers to further examine the robustness of student learning (with the applied flag activity), which showed evidence that students could transfer their understanding to a related setting.

5 Retrospective analysis

Analysis conducted *during* a teaching experiment focuses on supporting the participants' learning. *Retrospective* analysis, by contrast, allows researchers to consider the design research study in a broader theoretical context of the larger (but still domain-specific) issues of the overall study.

> The situated nature of retrospective analyses is a strength of the methodology, given the overall goal of engineering new forms of learning and the tendency of "high" theory to pass over what may be important details in an effort to paint phenomena in uniform terms. In particular, because the resulting accounts of learning are tied to the means by which it was generated, the design team is always in a position to develop testable conjectures about how those means of support and, thus, the instructional design might be improved. (Cobb et al. 2003, p. 13)

The team considered the original research questions and sub-questions concerning the effects of the curricular materials, the instructional design and the use of the IDAS. These were juxtaposed to the detailed evidence concerning the two conjectures about sharing and naming collections and sharing wholes. The team had expected numerous

---

[5] Prior research has shown that individualized instruction has uneven and limited benefits (Erlwanger 1975).

Table 1 Packet 1 diagnostic assessment results (collections), pre- and post-assessment

	6 Among 2	8 Among 2	10 Among 2	14 Among 2	16 Among 2	6 Among 3	9 Among 3	12 Among 3	15 Among 3
Admin 1 (%)	100	100	66.7	100	63	100	0	100	100
n	3	6	3	4	8	2	1	5	5
Admin 2 (%)	75	100	100	85.7	100	100	100	50	66.7
n	4	3	2	7	2	1	1	2	3

	18 Among 3	8 Among 4	12 Among 4	16 Among 4	10 Among 5	15 Among 5	12 Among 6	18 Among 6
Admin 1 (%)	100	50	66.7	60	100	50	100	100
n	1	2	6	10	4	6	4	2
Admin 2 (%)	66.7	50	100	66.7	80	66.7	60	50
n	3	2	2	3	5	3	5	6

Whole-class percent correct for item type (number of items among number of people *p*)

n number of students who received item

challenges—technical, practical, and cognitive—because it was the first time the IDAS software and the curricular materials had been implemented. The retrospective provided an opportunity to consider these components as part of a single learning ecology.

5.1 Findings on and implications from the conjectures

To begin, we examined the data collected by the IDAS system for sharing collections and, later, the data for sharing wholes.

5.1.1 Examining the data from Conjecture 1

Table 1 shows the class's overall performance on collections tasks (packet 1 items), listed by task parameters for two administrations (day 1 and day 3). The task parameters (number of objects, number of people sharing) are ordered by number of people sharing, from low to high, and secondarily by the number of objects, from low to high. The *n* values vary because items were randomly assigned per student for each diagnostic by sampling of parameters by IGE.

Further analysis showed that the average performance on a set of items declined as *p* increased, with the exception of sharing by 5's. It should be noted that unfamiliarity with the IDAS software, especially with touch screens, may have produced some unreliable results. Across all students in the study despite small *n*'s, modest gains were seen on 2's 4's and 5's, with losses on 3's and 6's. However, because items were randomly assigned, individuals' score comparisons across administrations do not permit conclusions regarding improvements (Table 2).

Table 2 Packet 1 diagnostic assessment results (collections)

p	2	3	4	5	6
Admin 1 (%)	83.3	92.9	61.1	70.0	100
n	24	14	18	10	2
Admin 2 (%)	88.9	70.0	71.4	75.0	50.0
n	18	10	7	8	6
All items (%)	85.7	83.3	64.0	72.2	62.5
Total *n*	42	24	25	18	8

Whole-class percent correct for each item type, representing the average of scores on all items involving sharing among *p* persons

n number of students who received item

Table 3 displays class-level scoring for the packet 1 naming questions (naming questions were administered only if a student correctly completed the equipartitioning task). Of those items given to students on the first administration, a majority of the count questions were answered correctly (58 % correct) but only 9 % of the relational and reassembly questions. All naming question scores showed credible progress on the second administration: 73 % of the count questions, 27 % of the relational naming question and 32 % of the reassembly questions were answered correctly.

The results in Table 3 call into question students' performance on Conjecture 1. Students' difficulties with the naming questions (based on classroom discussion, LPPSync results, and students' own verbal feedback) confirmed that relational naming and reassembly were challenging and novel, and the students showed modest improvement.

That is why the team concluded that the curriculum and the diagnostic assessment require revision to adequately support student learning of relational naming. LPPSync design had not anticipated the language the students seemed to prefer ("3 out of 12"). The design research study

Table 3 Packet 1 (collections) naming and reassembly item responses

Item type	Administration 1 (*n* = 33)			Administration 2 (*n* = 22)		
	% C	% I	% NR	% C	% I	% NR
Counts	58	30	12	73	18	9
Relational	9	82	9	27	41	32
Reassembly	9	45	45	32	23	45

C correct, *I* incorrect, *NR* no response, *n* number of students receiving an item

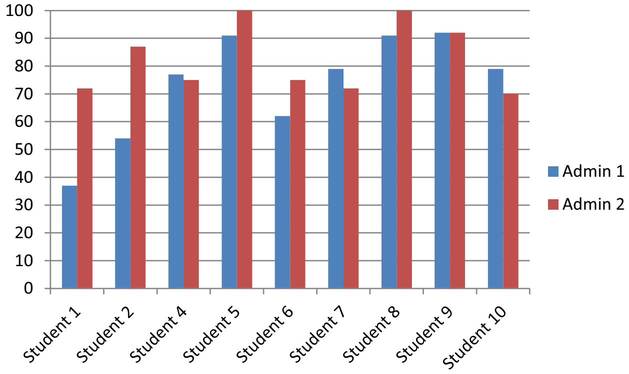

Fig. 5 Student proficiency scores for two administrations of packet 2 (wholes) diagnostic. Only 10 students took both assessments

provided direction for design revisions by showing that relational naming can be scaffolded with proto-relational naming, along with a discussion of reassembly. Students may be better able to secure their understanding of relational naming of collections by examining the relationship between sharing a collection and sharing a single whole only after equipartitioning and naming the shares of single wholes. After the study, the team set about designing revisions to place only rudimentary naming questions with the equipartitioning tasks in packets 1 and 2, to develop a new packet that focuses on relational naming and reassembly for both collections and wholes, and to redesign the LPP-Sync naming question interface.

5.1.2 Examining the data on Conjecture 2

Figure 5 displays individual student scores for the two packet 2 diagnostic administrations on sharing wholes (day 4 and day 6)[6]. Many students' scores improved from the first to the second assessment, but in light of robust class discussions around the criteria for equipartitioning, *any* decrease surprised the team. The subsequent paper and pencil formative assessment and the flag activity (responses to the diagnostic data) strengthened the children's use of the criteria to justify their equipartitioning task solutions.

5.2 Findings, implications, and theory building on the research questions

Having discussed the specific conjectures, we return to examination of the overall research questions and sub-questions. Broader than the conjectures, these involve overall questions about learning that devolve from the interventions and the sub-questions on (1) strategies, naming, justifying and reassembly and (2) the use of the IDAS.

5.2.1 Sub-question 1 on equipartitioning strategies, naming, justifying and reassembly

The children in this study began with insufficient experience in equipartitioning, despite the fact that most of them were entering fourth grade and should have had an introduction to fractions in third grade. The study supports the claim that direct experience with equipartitioning collections and wholes needs more emphasis in early grades. It suggests that children at these ages are highly capable of learning to equipartition collections and wholes, but need the complete set of carefully sequenced of tasks to be fully successful. The LT specifies sharing collections as the first level and confirms that students' experiences with dealing allow them to be reasonably successful with sharing collections and providing answers to the question "how many objects are in a single person's share?" Explicit discussion about ways other than counting to confirm and justify their answers is valuable (i.e. showing one-to-one correspondence and ensuring the completion of all rounds of dealing). In sharing a rectangular whole, students needed to engage with the fact that the number of parallel lines drawn must be one fewer than the number of shares desired. In the circular whole, a direct discussion of the misconception that "equally spaced vertical lines result in fair shares" needs to be made explicit and examined. In sharing circles, students quickly master sharing for the powers of 2, but creating an odd number of shares required a more direct intervention for conceptualizing the necessary radial cuts. Finally, the students benefited from articulating the three criteria for successful equipartitioning, and incorporated the criteria language into their justifications. Overall, the study supports the view that the first two levels of the trajectory are essential as a means of helping students to learn how to

---

[6] Figure 5 shows correct scores for equipartitioning tasks, based on replaying recorded student on-screen actions, circumventing the software scoring defect (see footnote 4).

carry out the operation of equipartitioning for collections, circles, and rectangles. Further research could examine the degree to which, if any, the results generalize to other shapes. As reported above, students also need careful scaffolding to learn how to name the results relationally and to describe the process of reassembly as times as many.

The children's analysis of the flags provided evidence that they were developing not only a means and a language to apply equipartitioning, but that some were also able to extend its principles to make other arguments about relative size of regions. Though tentative, the results suggest the value of forming a stronger foundation in the activities of equipartitioning in preparation for the introduction to fractions.

5.2.2 Sub-question 2 on the effects of the IDAS

Using any new technologies in the context of classroom practice, especially wireless connectivity, involves significant challenges. For its first deployment in real-time use, the LPPSync system technology held up reasonably well. Students were able to take the diagnostics and input their responses. The data were successfully stored and retrieved, and reports were generated. Children liked using the iPads, found direct action on the screen, to share and equipartition, interesting and worthwhile, and found it easy to collaborate in the practice mode. When they received feedback in real time, the students were delighted and engaged.

A diagnostic process can be evaluated on several bases. It should provide:

- valid indicators of domain-appropriate cognitive function,
- valid targets for diagnosing acceptable levels of function,
- appropriate protocol for assessments and for teacher and student feedback,
- accurate, useable reports,
- support for personalization of treatment,
- indication or a related source for an appropriate course of action,
- support for coherence of instruction and learning.

Most these criteria were met at least rudimentarily during the study. The IDAS contributed to a data-rich learning environment in which students and teachers worked in an informed and focused manner to improve learning. It promoted learning focus on student learning of LT proficiencies, revealing which items and concepts within the EPLT students seemed proficient with, and with which items, parameters, or concepts individual students or the entire class had difficulty. These data allowed the teacher to provide more adept, finely-tuned, and differentiated instructional moves and planning. The IDAS also contributed to the conduct of the study by providing data upon which to periodically test the conjectures. Further design and interface refinement could enable the IDAS to provide individual students with explicit feedback on what kinds of problems they have mastered and which ones still need to be improved (done by hand by research team during this study).

Secondarily, the IDAS contributed to a climate of higher student expectations for retaining information and working steadily for improvement. Regular personal results gave students a sense that their own performance matters and is supported, which in turn reinforces their intellectual growth, sense of responsibility for their own learning, and the sense that the class as a whole is moving towards increased proficiency. These are all features of strong formative assessment (Heritage 2010).

Refinement and eventual large-scale use of such software innovation requires ongoing attention to design, function, and implementation, to improve the interplay of domain-specific cognitive issues *and* user experience. We gained numerous insights regarding the IDAS design throughout the study. Several particular issues—calibration of software scoring algorithms to instructional needs, inadequate interface design for the naming questions and response options—surfaced as major concerns. The results and insights gained from the design research study will lead to major improvements in all the components of the learning ecology.

5.3 Student response to the learning ecology

Design research studies allow researchers to examine the effects of the designs, not only in terms of direct cognitive outcomes, but also with regard to the impacts on students as members of a learning community—including their uptake of mathematically productive norms of behavior (Yackel & Cobb 1996). Team members repeatedly commented on improvements they observed in participants' overall approach to learning, self-confidence to learn the material, and persistence, systematicity, and information retention. Students reported (to evaluators) that participating in the design research study helped them help each other (i.e. to collaborate) solve problems, and that the tablets (which were new to the students) were fun to use and made it easier to stay focused on solving the math problems. Students also spoke of their determination to understand and improve their mathematics learning, and stated emphatically that they were better students because of the two-week class. Future research studies will incorporate more information on these critical elements of achievement motivation (Dweck and Leggett 1988; Pintrich and Schunk 2002).

6 Design research study contributions to theory

Design research study intends to make important contributions to theories of learning and teaching, due in part to its engineering perspective and in part to the careful observation of the interactions of the components. This particular study was designed to explore the interactions among a set of curriculum materials, a digital diagnostic assessment system (both based on an underlying cognitive framework), and an instructional setting with high-needs students.

The intervention was novel (new materials, new technology), and incorporated the challenge of working with high-needs children who were unaccustomed to collaborative instructional settings. Many factors were in play simultaneously—helping students learn to manage the setting and revise expectations and behaviors, implementing new, previously unpiloted curriculum materials, and deploying a new technology dependent on reliable connectivity to implement a wireless infrastructure. Despite all of this instability and implied high risk, the research team found considerable overall progress on the EPLT as a backbone of instructional change. Students learned and improved their understanding of equipartitioning—gradually but steadily.

A key understanding about equipartitioning was the need for extensive individual experience at the level of the concrete operation of equipartitioning, both of collections and of single wholes. The curriculum materials systematically varied both the sizes of the collections and the numbers of people sharing, so that students experienced each different type of challenge. While class discussion and collaborations can surface and spread strategies, unless each individual student experiences the direct need for instance, to share for an odd number instead of an even number of persons, he or she is unlikely to be successful. This reinforces the value of a diagnostic tool designed to periodically check individually for these proficiencies. Successful behavior among the class as a whole does not imply each student is able to accomplish the task. Our diagnostic system demonstrated that often teachers, and researchers, overestimate student success.

A second key understanding, generated from student work on sharing a whole, was the essential role class discourse provided in students articulating the set of three criteria as an explicit means of checking and justifying equipartitioning. While some students intuitively recognized the need for these, systematically identifying and applying them to justify a correct answer evolved to be a classroom norm.

A third recognition about learning concerned the differences in the way in which the activity of naming and discussing the results of fair sharing and the results of reassembly evolved. These elements of the learning trajectory required much more scaffolding and reinforcement than initially anticipated. While they can be introduced as students become proficient at the earlier levels of the trajectory, deep mastery required them to apply and contrast the naming conventions in both sharing collections and wholes. While both relational naming and reassembly (“times as many/much”) had roots in everyday activity around halving and doubling, extending those meanings to $1/n$th or n times as many/much required practice and gradual reinforcement across cases. This insight made us realize that including naming and reassembly in the first diagnostic assessment was not viable or effective; while it should remain in the packets as students learn to carry out equipartitioning, rapid mastery should not be expected.

Finally, the study showed us that forming relationships among the ideas, even at the lowest levels of the trajectory (such as how $1/n$th of a collection is similar to and different from $1/n$th of a whole, or how “$1/n$th of” as relational naming anticipating division and ratio, and “times as many/much” anticipating multiplication as an inverse operation), takes a different kind of interaction, multiple passes, and discussion with students about what they expect to learn.

In these ways, one sees how a design research study contributes to theory of understanding, and plays a critical role in the development of theory about domain-specific learning. Learning is not simply based on the discovery of students’ natural inclinations, but are instead grounded in the design of sequences that both leverage those natural tendencies and instructionally promote strong pathways to sophisticated mathematics. While this particular research study illustrated these key conclusions with respect to the equipartitioning learning trajectory, related curriculum materials, and instruction, the insights gained can be applied more broadly.

Acknowledgments This material is based upon work supported by the National Science Foundation (DRL-0758151) and Qualcomm. Opinions, findings, conclusions, or recommendations are those of the authors and do not necessarily reflect views of the funders.

References

Artigue, M. (2008). Didactical design in mathematics education. In C. Winslow (Ed.), *Nordic research in mathematics education: proceedings from NORMA08 in Copenhagen, April 21–25, 2008* (pp. 7–16). Rotterdam: Sense Publishers.

Barrett, J. E., & Battista, M. T. (2014). A case study of different learning trajectories for length measurement. In A. P. Maloney, J. Confrey, & K. H. Nguyen (Eds.), *Learning over time: learning trajectories in mathematics education* (pp. 97–124). Charlotte: Information Age.

Black, P., & Wiliam, D. (1998). Inside the black box: raising standards through classroom assessment. *The Phi Beta Kappan, 80*(2), 139–148.

Brown, A. L. (1992). Design experiments: theoretical and methodological challenges in creating complex interventions in classroom settings. *The Journal of the Learning Sciences, 2*(2), 141–178.

Clements, D., & Sarama, J. (2009). *Learning and teaching early math: the learning trajectories approach*. New York: Routledge.

Clements, D. H., Wilson, D. C., & Sarama, J. (2004). Young children's composition of geometric figures: a learning trajectory. *Mathematical Thinking and Learning, 6*, 163–184.

Cobb, P., Confrey, J., diSessa, A. A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.

Collins, A. (1992). Toward a design science of education. In E. Scanlon & T. O'Shea (Eds.), *New directions in educational technology* (pp. 15–22). Berlin: Springer.

Confrey, J. (1994). Splitting, similarity, and rate of change: A new approach to multiplication and exponential functions. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 293–332). Albany: State University of New York Press.

Confrey, J. (2006). The evolution of design research studies as methodology. In K. R. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 135–152). New York: Cambridge University Press.

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

Confrey, J., & Lachance, A. (2000). Transformative teaching experiments through conjecture-driven research design. In A. Kelly & R. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 231–265). Mahwah: Erlbaum.

Confrey, J., & Maloney, A. P. (2012). Next generation digital classroom assessment based on learning trajectories in mathematics. In: C. Dede, & J. Richards (Eds.), *Digital teaching platforms: customizing classroom learning for each student* (pp. 134–152). New York: Teachers College Press.

Confrey, J., Maloney, A. P., & Corley, A. K. (2014a). Learning trajectories: a framework for connecting standards with curriculum. *ZDM–The International Journal on Mathematics Education 46*(5), 719–733.

Confrey, J., Maloney, A. P., Nguyen, K.H., Lee, K., Panorkou, N., Corley, A. K., & Neal, L. (2011b). *Equipartitioning curriculum packets 1–7. A learning trajectory-based mathematics curriculum* (unpublished).

Confrey, J., Maloney, A. P., Nguyen, K. H., 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). PME: Thessaloniki, Greece.

Confrey, J., Maloney, A. P., Nguyen, K. H., & Rupp, A. A. (2014b). Equipartitioning, a foundation for rational number reasoning: Elucidation of a learning trajectory. In A. P. Maloney, J. Confrey, & K. H. Nguyen (Eds.), *Learning over time: learning trajectories in mathematics education* (pp. 61–96). Charlotte: Information Age.

Confrey, J., Maloney, A. P., Wilson, P. H., & Nguyen, K. H. (2010). *Understanding over time: the cognitive underpinnings of learning trajectories.* Paper presented at the Annual Meeting of the American Education Research Association, Denver, CO.

Design Based Research Collective. (2003). Design-based research: an emerging paradigm for educational inquiry. *Educational Researcher, 32*(1), 5–8.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*(2), 256–273.

Erlwanger, S. H. (1975). Case studies of children's conceptions of mathematics: I. *Journal of Children's Mathematical Behavior, 1*(3), 157–283.

Gravemeijer, K. (1994). Educational development and developmental research. *Journal for Research in Mathematics Education, 25*(5), 443–471.

Gravemeijer, K. P. E., & Cobb, P. (2006). Design research from a learning design perspective. In J. Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 45–85). London: Taylor Francis Group.

Heritage, M. (2010). *Formative assessment and next-generation assessment systems: are we losing an opportunity?*. Washington, DC: Council of Chief State School Officers.

Lehrer, R., Kim, M.-J., Ayers, E., & Wilson, M. (2014). Toward establishing a learning progression to support the development of statistical reasoning. In A. P. Maloney, J. Confrey, & K. H. Nguyen (Eds.), *Learning over time: learning trajectories in mathematics education* (pp. 31–59). Charlotte: Information Age.

Piaget, J. (1970). *Genetic epistemology*. New York: W.W. Norton & Company Inc.

Pintrich, P. R., & Schunk, D. (2002). *Motivation in education: theory, research and applications*. Upper Saddle River: Prentice-Hall.

Plomp, T. & Nieveen, N. (Eds.) (2013). *Educational design research: illustrative cases.* Enschede: SLO, Netherlands Institute for Curriculum Development. (http://international.slo.nl/publications/edr/). Accessed 25 May 2014.

Popham, W. J. (2008). *Transformative assessment*. Alexandria: Association for Supervision and Curriculum Developoment.

Pothier, Y., & Sawada, D. (1983). Partitioning: the emergence of rational number ideas in young children. *Journal for Research in Mathematics Education, 14*(5), 307–317.

Streefland, L. (1984). Search for the roots of ratio: some thoughts on the long term learning process (towards…a theory): part I: reflections on a teaching experiment. *Educational Studies in Mathematics, 15*(4), 327–348.

van den Akker, J., Gravemeijer, K., Mckenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research: the design, development and evaluation of programs, processes and products*. London: Routledge.

van Galen, F., Feijs, E., Figueiredo, N., Gravemeijer, K., van Herpen, E., & Keijzer, R. (2008). *Fractions, percentages, decimals and proportions: a learning-teaching trajectory for grades 4, 5 and 6*. Rotterdam: Sense Publishers.

Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458–477.