From handheld collaborative tool to effective classroom module: Embedding CSCL in a broader design framework

Jeremy Roschelle a, Ken Rafanan a, Gucci Estrella a, Miguel Nussbaum b,*, Susana Claro b

a SRI International, Menlo Park, CA 94025, United States
b Pontificia Universidad Católica de Chile, Santiago, Chile

ARTICLE INFO

Article history:
Received 19 October 2009
Received in revised form
20 April 2010
Accepted 21 April 2010

Keywords:
Computer-supported collaborative learning
Framework for CSCL practice
Mathematics learning
Handheld-supported learning
Formative assessment
Design-based research

ABSTRACT

The field of Computer Supported Collaborative Learning (CSCL) includes designers who emphasize effectiveness, measured via experiments, as well as designers who emphasize context and conduct qualitative research on teaching and learning practices. We conjectured that these two different emphases could be fruitful combined in a research and development process aimed at producing effective CSCL practices. We explored this possibility in a project that adapted a CSCL tool from Chile to serve as the basis of an effective 3-week classroom module for primary school mathematics in the United States. To go from tool to module, we addressed curricular fit, training materials, pedagogical guidance, formative and summative assessments, and logistical support. In conducting the project, we found that effectiveness and contextual research could be conducted simultaneously and yielded complementary insight to this design process, which enabled our project to rapidly move from the base tool towards complete classroom modules. An experiment we conducted after our design iterations showed that students who used the modules learned more about the target content, fractions. A retrospective analysis of our design process suggests that the Integrative Learning Design framework is useful for organizing the complementary components of effectiveness and contextual research in our design process.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

The field of Computer Supported Collaborative Learning (CSCL) focuses on the enhancement of student learning through students’ use of technologies that have been specifically design to enhance students’ collaboration in educational activities (Crook, 1994; Dillenbourg, 1999). When students have the opportunity to work in small groups, they can contribute to a common understanding as well as develop verbal and social abilities (Nussbaum et al., 2009). Although researchers have reported results from the use of CSCL tools (as discussed in more detail shortly), little influence has been observed on actual school practice outside the context of research investigations. Brown (1994) argues that school practices are still influenced by “outmoded theories of learning” because “what new theories ask is so hard.” However, with the increasing expansion of technology in the classroom (laptops, netbooks, PDAs, phones), there is hope that new practices can become feasible for broader adoption.

In the CSCL field, an emphasis is on what might be termed CSCL effectiveness—comparing CSCL approaches with non-CSCL approaches to the same subject matter. Random assignment experiments are a powerful methodology for such experiments, and an increasing number of experiments are being conducted in CSCL, particularly in validating CSCL “scripts” (e.g., Schoonenboom, 2008). From a design point of view, however, CSCL has tended to emphasize contextually rich research traditions, for example to investigate the construction of new practices of teaching and learning from the contributions of multiple social actors in the classroom. Design-based research methodologies, as described by the Design-Based Research Collective (2003), suggest blending empirical educational research with theory-driven design of learning environments to better understand educational technologies in context.

A benefit of design-based research is the ability to collect and analyze rich data on many factors simultaneously and to use this rich data to iteratively improve a design more rapidly than might be accomplished through systematic experimentation on each individual factor.
Thus, design-based research does not necessarily include a causal analysis of effectiveness and thus does not ordinarily require randomized trials. Indeed, the Design-Based Research Collective does not recommend randomized trials at all. In its view, randomized trials may “hinder innovation studies by prematurely judging efficacy” and may fail to account for the phenomena that “educational research most needs to account for in order to have application to educational practice.” The collective is referring to the many detailed decisions about the design and implementation process (2003). Yet other design-based research theorists such as Bannan-Ritland (2003) see randomize trials as “stage-sensitive” and thus best left to the judgment of the researchers to include at any appropriate time. Other researchers (Shavelson, Phillips, Towne, & Feuer, 2003) consider randomized trials to be essential.

As Brown (1994) argues, laboratory-only results often cannot be reproduced in school settings. Therefore, instructional design must rise from both theory and practice in real context, where “context” is part of the analysis. Attending to the context helps produce better understanding of the intervention and improves the theories that argue for it. In design-based research methods, both “context” and “intervention” inform how to implement an instructional design (Design-Based Research Collective, 2003). Unlike descriptive theories (such as learning theory), design theory (Bannan-Ritland, 2003; Design-Based Research Collective, 2003; Reigeluth & Frick, 1999) suits settings in which some aspects of the intervention can emerge from an interaction between design goals and the school context (Reigeluth & Frick, 1999).

For CSCL designs to grow in influence by establishing their effectiveness, we saw a need to combine contextual and effectiveness perspectives to generate more research-based knowledge about effective CSCL practices. Such a conjoint perspective might not only prove effectiveness, but also pay attention to what prevents schools from implementing changes (such as CSCL) in their practice despite their proved benefits, and thus contribute to eventual adoption. This paper reports on a project called TechPALS that provided an opportunity to explore how the design of effective CSCL practices could be enhanced by combining effectiveness and contextual perspectives.

The TechPALS project investigated whether an existing face-to-face CSCL tool from Chile called Eduinnova might serve as the basis of an effective CSCL practice in U.S. primary school classrooms. (TechPALS was chosen as a name to stand for “technology for peer assisted learning” and to convey technology an emphasis on sociality through the connotation of a “pal.”). Eduinnova is a suite of software activities and a database of content for wireless handheld platforms that support a collaborative methodology. Although these activities were not specifically mathematical, they could be adapted to mathematics tasks. To expand from Eduinnova to a classroom module, we integrated it with a portion of the U.S. mathematics curriculum, specified how it fit into U.S. instructional practices, designed training materials and supporting classroom routines, incorporated formative assessment practices, and identified appropriate summative assessments. To evaluate whether the resulting classroom module, called TechPALS, was effective and to identify the improvements it would need, we conducted a series of research cycles, including randomized experiments in three schools in our pilot year and three more schools in the subsequent experimental year. As we describe here, the pilot year yielded mixed results, which led to significant design iteration starting a new cycle. Some of the improvements were in the tool itself, but many were in the supporting framework. After applying these improvements, we obtained statistically significant effects in our school experiments: Students who used TechPALS learned more.

In this paper, we use these experimental results and information gathered from observations to provide a context for discussing the broader design framework that was needed to transform a tool to an effective classroom module. To organize our observations, we retroactively apply an existing framework for design-based research called Integrative Learning Design (ILD). In describing ILD, Bannan-Ritland (2003) distinguishes four stages of research and development (Fig. 1): (1) Informed Exploration, studying needs, theory, and audience; (2) Enactment, designing the prototype or intervention and implementing and testing it in as many cycles as necessary; (3) Evaluation: Local Impact, evaluating the local impact of the intervention on its “clients” and going back to the enactment phase if necessary; and (4) Evaluation: Broader Impact, paying attention not only to publication and dissemination of findings, but also to concerns related to the adoption of the intervention by a broader audience. Although the ILD framework was not used during the TechPALS project, the project fits the framework. In each phase of the evaluation, both field observations and quantitative results were analyzed. As in the ILD framework, when either aspect of the analysis showed need for improvement, we began a new cycle, going back to the Enactment stage to improve and enrich the instructional design.

After a brief literature review and presentation of the Eduinnova approach that provided the base technology and activities for our efforts, we describe two design cycles that advanced from the base technology towards an effective CSCL practice:

- **Year 1, From CSCL tool to pilot classroom module.** This cycle started at the ILD stage of Informed Exploration and proceeded to the first Enactment stage and Evaluation of local impact. The Enactment stage included the design process of turning the tool into a classroom module and implementing the pilot. Although the pilot included a randomized experiment, the results of the evaluation of the pilot’s local impact are mainly qualitative, showing that the design was not mature enough yet.

- **Year 2, From pilot classroom module to implemented classroom module.** In this cycle, we revisited the Enactment stage, improving the classroom module and implementation design using information from the first cycle. The implementation included a randomized experiment, followed by a new phase of Evaluation of Local Impact. For the evaluation in this cycle, we were able to take valuable information from both the summative assessment of the randomized experiment and the observations of the implementation in school contexts.

---

![Fig. 1. ILD framework for design-based research.](https://example.com/fig1.png)
2. Literature review: handheld CSCL.

As used in educational research, the terms “cooperative learning” and “collaborative learning” convey both similarities and differences in objectives and methodologies (Mueller & Fleming, 2001; Panitz, 1999; Teasley & Roschelle, 1993). For the present investigation, the similarities are more useful than the differences. The objective of cooperative learning is to attain certain instructional goals through work in small groups. Characteristic of group work methodologies is the definition of the roles the group participants must take on and the tasks they must carry out in pursuit of these goals (Cooper & Others, 1991; Ellison, Boykin, Tyler, & Dillihunt, 2005). The key to this type of learning is the appropriate distribution of tasks and responsibilities for achieving the goals set by the group’s instructor (Mueller & Fleming, 2001; Panitz, 1999). Adams and Hamm (1996) and Dillenbourg (1999) have established five factors that contribute to effective collaborative learning:

1. Individual responsibility. Each member is responsible for his or her own work, role, and efforts to learn within the group (individual rules and roles) (Slavin, 1996).
2. Mutual support. In addition to being responsible for his or her own learning, each member is responsible for helping to teach other members of the group through the frequent exercise of social skills during group interactions.
3. Positive interdependence. The main aim of the activity is the group goal. Success is therefore achieved only when all team members have reached their individual goals.
4. Face-to-face social interaction. Decision making must involve discussion among the members of the group. Productivity is therefore affected by the group’s ability to efficiently exchange opinions and make compromises to build a consensus answer.
5. Formation of small groups. Discussion, social interactions, and consensus building can be achieved only in small groups of three to five members (Adams & Hamm, 1996).

Building on this, the objective of collaborative learning is to promote positive interdependence and individual responsibility among group participants to reach a goal they all share (Cooper & Others, 1991; Panitz, 1999). This requires methodologies that enable the participants to comprehend the goal of an activity through interaction and negotiation of their individual perspectives (Mueller & Fleming, 2001; Panitz, 1999). Collaborative learning seeks to create an emotional environment that favors dialogue between participants so they can construct a shared learning experience (Panitz, 1999).

As Zurita and Nussbaum (2004b) indicate, research has shown the following benefits of collaborative learning for children:

- Seeing classmates as sources of knowledge and help rather than as competition
- Developing common values, social skills, and ability to participate in teamwork
- Emphasizing thought and understanding instead of memorization
- Articulating ideas and navigating among others’ ideas
- Enjoying and learning more by being active participants in their learning.

3. The eduinnova approach to classroom CSCL.

In the fields of CSCL and mobile learning, investigators have developed a number of approaches for using handheld devices to support collaboration among students (Chan, Roschelle, Hsi, Kinshuk Sharples, Brown, et al., 2006). Many of these approaches emphasize applications in museums (Yatanai, Sugimoto, & Kusunoki, 2004) or in the outdoors (Tan, Liu, & Chang, 2007), where the mobility of handhelds is required because of the settings. Their low cost and ease of integration into everyday classroom routines makes handhelds also attractive for in-school uses (Roschelle & Pea, 2002). We begin by describing the Eduinnova approach to in-school uses of CSCL because this provided the base design for our research and development process.

Focusing on in-school uses, Nussbaum and colleagues designed software that runs on low-cost mobile devices to support collaborative activities among students working in small groups on typical school subject matter (Cortez, Nussbaum, Rodriguez, Lopez, & Rosas, 2005; Zurita & Nussbaum, 2004a). Zurita and Nussbaum’s approach recognizes two networks: the social network where group mates interact verbally and the technological network that transparently supports the social network activities by coordinating and synchronizing activity states and mediating the activities and the social interaction of the participants (Zurita & Nussbaum, 2004a). The software and the related activities that implement this approach are called Eduinnova. After successful initial trials in Chile, Nussbaum sought collaborators in other countries who could leverage Eduinnova in further research and development. SRI International decided to test it for teaching fractions to fourth-grade (age-9) students in U.S. primary schools.

The Eduinnova software is intended to support both collaborative learning and formative assessment by managing the work of assigning roles in collaborative tasks to students, the software can make implementing CSCL in the classroom easier. Further, by providing rapid feedback to students and teachers in a relevant and comprehensible format, Eduinnova can make implementing formative assessment easier. “Immediate feedback” denotes rapid, useful feedback to students and teachers and is related to the broader notion of “formative assessment.” In a meta-analysis of 58 studies, Bangert-Drowns et al.’s (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991) found a modest overall positive effect (+0.26) for feedback on student achievement. In another meta-analysis, Kluger and deNisi (1996) found higher effects when students were given feedback on the correctness of their solution methods and on their improvement from earlier trials and when they were using computers. Finally, other individual investigators have found that some of the most effective forms of feedback (1) guide improvement on product as a student is making it or (2) guide teachers to adjust students’ instruction (Butler & Winne, 1995). In fact, researchers (Fuchs & Fuchs, 1986) found large positive effects for achievement when teachers were required to alter their educational program for students based upon the feedback.

Four Eduinnova activities fit our target domain of fourth-grade fractions. In the Consensus activity, each student in a group of three receives the same multiple choice question at the same time (Fig. 2, Activity 1). Each student enters an answer independently, but the
system requires that the students agree on the answer and provides feedback only at the group level. Once students agree, the software tells students either that everyone chose the correct answer or that at least one student (not identified) was wrong (immediate feedback). The group may not go to the next problem until all students have answered correctly. Since the software does not tell students which students were right or wrong, discussion among students is encouraged.

In the Exchange activity, each student receives two representations of a fraction (Fig. 2, Activity 2) with the goal of matching the representations on his or her screen. This is achieved if the representations depict equivalent fractions. Students exchange representations within their group until they think they have a match; the software gives positive feedback only if all the matches are correct. After three failed attempts, the software indicates the correct answer.

In the Ordering activity, each student in the group of three receives a unique fraction between 0 and 1. As a group, the students must input the fractions in a sequence of ascending order (Fig. 2, Activity 3). Each student must place his or her fraction at the right point in the sequence. Once all the group members have placed their fractions, the system evaluates the submitted sequence. If the group has submitted an incorrect sequence, the system gives it another try. After three tries, the software displays the correct answer.

In the Aiming Between activity, the group members receive the same representation of a number line (between 0 and 1) with a target interval highlighted, as shown in Fig. 2 (Activity 4). Each student proposes independently a fraction that should be within the target interval. Once all the group members have submitted their fractions, the system asks them to evaluate each of the three answers (correct/incorrect). Once the three answers have been evaluated in the same way, the system indicates whether the consensus evaluation is correct. If no answer is correct, the students repeat the evaluation process. After three incorrect evaluations, color-coded arrows are displayed on the number line corresponding to the three fractions along with a list of the submitted answers.

Fig. 2 displays examples of the screens for each of these activities. The screens for Activity 1, Consensus, show each of the three group members’ selected answer. The system has responded, “You all have to agree!” because the students did not all submit the same choice. The
screens for Activity 2, Exchange, show the start of the exercise when each of the students has received a pair of fractions in different representations. The students must decide whether to trade the lower representation or keep it if it matches the upper one. The screens for Activity 3, Ordering, show the different numbers assigned to each member (on the lower row) and the common area (a staircase), where each of the students must place the fraction that he or she received. One student has placed the first number (the orange $\frac{1}{4}$) because it is the lowest. The students must decide who should go next. For Activity 4, Aiming Between, we see the “fraction construction screen,” which displays the number line and the target that the students must aim for. Each student is creating his or her own fraction to submit to the group.

Across all activities, the teacher receives feedback on how the students are doing. The feedback is organized as a simple grid of groups (rows) by problems (columns), as displayed in Fig. 3. A cell in the grid is green if the group gets that problem right on the first try, yellow if the group gets the problem right on a later try, and red if the group exceeds the number of allowed trials. By scanning the grid, a teacher can identify groups that are having trouble and provide assistance. Alternatively, the teacher can focus on a particular problem that requires additional explicit teaching. Thus, the teacher can enact formative assessment by adapting their instruction to fit emerging student needs.

4. Year 1: design and pilot

Bannan-Ritland’s ILD framework (2003) describes several research stages, each of which includes research questions, data, and methods and design of instruments and artifacts. The TechPALS initial design process involved both development work and pilot testing, both of which occurred in the first year of the project, the 2006–2007 school year. The development work entailed adapting Eduinnova activity structures to fourth-grade fractions content and designing training materials for students and teachers. After the ILD first stage, we studied the challenges that fourth-grade students face in learning fractions, the curricular structure of San Francisco Bay Area schools, and relevant cognitive and developmental theory. As Brown stated, “awareness of the deep principles underlying disciplinary understanding” is necessary for an improvement of instruction (1994). At the same time, we refined our understanding of how to apply collaborative learning theory to this challenge. Details are described in Roschelle et al. (in press). The pilot testing involved identifying or designing measurement instruments and conducting preliminary field tests in classrooms, tasks that correspond to the first cycle of an ILD Enactment phase. Our process for the development and pilot testing is described below.

4.1. Initial design of classroom modules

Our first decision in fitting the CSCL activity into a larger instructional framework was which phase of instruction we should target for collaborative learning. In broad terms, we saw instruction as composed of three phases: (1) teacher-led presentations (introduction of the math of the day) and discussion about collaborative skills, (2) student-centered practice, and (3) homework. We decided to deploy TechPALS collaborative learning as an alternative to individual student-centered practice.

To develop suitable activities for this instructional phase, we organized the content of fourth-grade fractions into two categories, concepts and procedures. We decided to focus on three important concepts of rational numbers: number, part-whole, and measurement. We organized our thinking about rational number procedures into three categories: operations (adding and subtracting), equivalence, and
ordering (including comparison). The design team then proposed four activities, each based on the Eduinnova activities described above, that would collectively cover the concepts and procedures (Table 1).

Understanding that the concepts and procedures were interconnected, we did not propose teaching them in isolation; rather, multiple activities enabled us to vary the emphasis and ensure coverage. The Consensus activity, because it is based on multiple-choice questions, could address all the concepts and procedures, but we thought some variety in activities would maintain a higher level of student engagement. We thought the Exchange activity would be particularly appropriate for the concepts and procedures relating to equivalent fractions because it requires students to match different representations of the same quantity. For example, the students might be challenged to match the fraction 1/4 to a pie divided into eight equal slices, two of which are shaded. In each group of three students, students would have three equivalent fractions to match, expressed in either the same or two different representations. The Aiming Between activity was chosen to focus on the concept of a fraction as a number on a number line, requiring students to construct fractions rather than choose among numbers already expressed as fractions. Finally, the Ordering activity introduced a new form of engagement because it required students to place their fraction in an ordering from smallest to largest. After the activities were determined, we developed databases of content for the individual items in each activity based on a detailed analysis of the content in the curriculum and research suggesting the kinds of problems that students would find to be difficult.

From prior research (e.g., Webb & Palincsar, 1996), we knew we would need to train students and teachers in collaborative learning behaviors. For this, we developed “The Cooperagent,” a teaching block that introduced and reinforced students' understanding of basic collaborative skills (such as asking for explanations and explaining why). It consisted of a short multimedia presentation and storybook about an agent who models collaborative learning behaviors. Consistent with the recommendations of Webb and Palincsar (1996), students were guided through The Cooperagent material to explain their answers and procedures and to ask for explanations, not just give “the answer.” This presentation occurred during the first session, as the first element of the intervention. Teachers received parallel training on their role in supporting collaborative learning.

### 4.2. Pilot testing

We initially tested TechPals in a real setting at a local San Francisco school, as a small pre-pilot test. We observed some technical issues, leading to technical improvements. On the basis of the pre-pilot, we also refined the evaluation instruments (such as the observation protocol) and the curricular content for the collaborative activities. We then planned the full pilot test.

The objective of the pilot test was to find weaknesses and strengths of both the classroom module and the implementation plan. Because a trade-off exists between refining the module for a specific context and generalizing the findings, we planned to conduct the pilot study in three different contexts. We tested our initial materials in three schools, each with different standings according to the California Academic Performance Index (API), used to rank California schools’ academic achievement from low to high:

1. A bilingual school in a major urban center, middle API
2. A school in an affluent suburban location, high API
3. A school in a relatively poor suburban location, low API.

In each case, the school used TechPALS for 2–3 weeks, but only schools 2 and 3 were teaching fractions while TechPALS was being implemented; school 1 taught fractions earlier in the year.

The following six key findings emerged from the observational and test score data we collected during the pilot testing:

1. TechPals approach did not work well in the poor suburban school because of severe behavioral problems in the classroom, despite the daily collaborative training given to students.
2. In the affluent suburban school, students did not learn very much because they were already too advanced.
3. The assessment we used, the Iowa Test of Basic Skills (ITBS), did not have enough fraction items to measure the learning we were observing. Further, the items that did exist on the ITBS were mostly procedural; the test did not pick up conceptual gains. We realized we needed to use a different end-of-unit test.
4. Ordering did not work well as a collaborative learning activity. This was the only activity in which students needed to wait their turn to act. In many cases, we observed one student directing the other students to press buttons in sequence; thus students did not do the activity collaboratively.
5. Technical problems (e.g. network downtime) distracted the teacher from guiding the activities which reduced students’ time on task.
6. Although collaborative behaviors were observed, we believed they could be more strongly encouraged while requiring less overall training time.

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Four activities covering target fractions concepts and procedures.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activity type</td>
<td>Concept</td>
</tr>
<tr>
<td></td>
<td>Number</td>
</tr>
<tr>
<td>Consensus</td>
<td>✔</td>
</tr>
<tr>
<td>Exchange</td>
<td></td>
</tr>
<tr>
<td>Aiming between</td>
<td></td>
</tr>
<tr>
<td>Ordering</td>
<td></td>
</tr>
</tbody>
</table>
5. Year 2, design refinement and randomized experiment

We started a new research cycle in our second year. In this cycle, we refined the design considerably, reshaping the classroom module as a result. The design refinements offer a look at the broader issues that must be tackled to integrate CSCL into ordinary classroom instruction. In addition, we conducted an experiment across three schools that demonstrated statistically significant results. Observations, however, suggested another phase of design would still be needed. As the Design-Based Research Collective would argue, design research provides a lens for understanding how to transform our theoretical claims into effective learning in educational settings by grounding the research “in the needs, constraints, and interactions of local practice” (2003). Following the research stages defined by the ILD framework, we returned to the Enactment phase after having evaluated the local impact of the pilot implementation.

5.1. Design refinement

In response to the the key findings, we made refinements to the research design. We discuss them, in order, from the broadly contextual to the narrowly technical.

First, we knew that we could not expect TechPALS to produce results in schools with severe behavioral difficulties because collaboration was unlikely when students were frequently misbehaving. Further, although we could have made TechPALS appropriate for advanced students if we had more time to produce adaptive databases of content, we realized the content would not be appropriate for affluent suburban students who were more than 1 year ahead of their lower income peers. Hence, we decided to target TechPALS to schools that were in the middle of the distribution incoming knowledge, not at the tails.

Second, to measure both procedural and conceptual gains, we switched from the commercial ITBS test to an established research-based test for primary school fraction content. Because the test was targeted to students who were 1 year older, we supplemented the test with some additional items for younger students.

Third, we tightened the fit of the TechPALS module to overall instruction. We analyzed the textbooks teachers would be using and specified an interleaving of teacher-centered presentations (without TechPALS) and student-centered practice with TechPALS, so that teachers would present the appropriate concepts and skills shortly before students would practice them. Further, whereas in the pilot a specially trained teacher aide introduced TechPALS, we provided teachers with more support so that they could be in charge of the activity. We did this because we had observed that students were less inclined to take the activities seriously when their ordinary classroom teacher was less involved.

Fourth, we sharpened The Cooperator agent collaboration training materials. To help students concentrate on the conceptual and procedural aspects of the math, we simplified the training to focus on the core collaborative learning skills of asking and answering two kinds of questions, how and why. With this focus, we were able to give students more examples of these skills. Further, to enhance the key elements needed for effective collaboration, we added a group challenge—the end-of-session test. In the group challenge, each student in the group answered a test question, working quietly and individually. Each student, however, received a score that was the sum of the correct answers in the group and these scores were publicly posted. Students quickly realized that a high score depended on their partners’ correct answers, so they were motivated to support learning for all members in their group during the practice sessions.

Finally, we dropped the Ordering activity (because it was susceptible to noncollaborative behaviors) and focused on only the remaining three activity types. In addition, we refined the technical design to reduce the network and login problems that had reduced time-on-task.

5.2. Experiment

Our research process included a comparative analysis of the effectiveness of TechPALS via a randomized experiment, which corresponds to the Local Evaluation stage of the ILD framework. We implemented the improved classroom module and gathered data to evaluate the local impact on the next ILD stage. We describe the experiment in brief here; a full description can be found in (Roschelle et al., submitted for publication), together with a description of the instruments and details of the findings.

We tested whether students assigned to the TechPALS intervention would outperform students assigned to work individually in a computer lab. We randomly assigned individual students to solve fractions problems during practice sessions using either TechPALS or the commercial software program iSucceed Math (formerly Larson Intermediate Math). This widely used commercial software provides students with a bank of practice items organized by topic.

We recruited two classrooms of fourth-grade students in each of three elementary schools that were in middle of the distribution of schools on California’s API. The school populations were approximately half Hispanic, 59% from families in poverty, and 45% English language learners. In the first half of each mathematics period, classroom teachers delivered their usual instruction to students. We arranged for half the students from one teacher’s classroom to exchange classrooms with half the students from a neighboring mathematics teacher for the portion of the period devoted to student-centered practice. Students used TechPALS for practice in one of the two newly mixed classrooms and iSucceed Math in the other. This design counterbalanced the effects of the different teachers across the two conditions. Further, the design ensured that students in both conditions spent the same amount of time practicing fractions with technology. Because we used random assignment to form the mixed classrooms, we have no reason to suspect any systematic bias due to the classroom of origin. Students were given a pretest on the first day of the experiment and an identical posttest on the last day of the experiment, spanning approximately 12 days of instruction and practice.

To examine group differences, we used a two-experimental-condition X three-school ANOVA with students’ gain score on the assessment as the outcome variable. We found a significant main effect of experimental condition (Fig. 4), with TechPALS students learning more [\( F(1,155) = 4.08, p < .05 \)]. In each school, the effect favored the TechPALS condition, but the effect size (Cohen’s d) ranged from 0.14 in school 2 and 0.17 in school 3 to 0.44 in school 1. Our observational data were consistent with our design premise: TechPALS would work by increasing student collaboration and improving feedback. Behaviors compatible with collaborative learning occurred significantly more frequently in the TechPALS condition. These include reading a problem aloud, asking a mathematical question, giving an explanation, making a collaborative move, directing a peer, and disagreeing with another student.
In analyzing the results of the last cycle, we found more than just an effect on learning. As Bannan-Ritland (2003) suggests, last stage must considers the consequences of the use of the product of the research for future practices or larger scales. In our consideration of potential consequences, we took into account not only the summative quantitative results, but also our own qualitative observations. Although the experiment produced an effect, our richer contextual observations also revealed some disappointments. In particular, we found that the textbooks the schools used presented concepts very poorly. Further, the textbooks presented conflicting procedures for slightly different types of problems. We also observed that teachers’ presentations of fractions topics tended to focus only on procedures. We observed students becoming confused while giving explanations to each other during the collaborative activities and inferred they could not be highly successful in collaborative learning if neither their textbook nor their teacher provided them with a conceptual basis for the mathematics.

5.3. Discussion

Although our initial research with TechPALS produced promising findings, additional work is needed before replicating it in similar settings. In particular, providing better curriculum materials are required for teachers and students and training is required for teachers so they can enact instruction that integrates concepts and procedures.

Our second step required extensive redesign. Most of the changes we made were broadly contextual, not tightly related to the collaborative activity. For example, we refined the specification of target schools, adopted a more appropriate assessment measure, and specified in more detail how the CSCL activity would fit into a larger unit of curriculum, instruction, and assessment. We added training material so classroom teachers could take more ownership of the CSCL portion of instruction. In particular, we tightened The Cooperagent training materials, added a group challenge to motivate students to help each other, and spent more time with teachers describing how their presentations should interleave with CSCL-based practice sessions. Some additional time was spent refining the Eduinnova tool and activities. The majority of our effort was contextual and not specific to the technology.

Whereas in the first year results had been mixed across schools, in the second year we measured learning gains with TechPALS in each of the three schools. Further, our observational measures showed that desirable collaborative behaviors were present in the TechPALS classrooms and not in the computer lab classrooms. This supports the conclusion that this iteration of TechPALS was an “effective CSCL practice.” Nonetheless, our results also suggested that further improvements would be possible.

6. Conclusion

Our conjecture was that design research can combine both effectiveness and contextual components. In reviewing our design process over two iterations, we found this to be true. With regard to effectiveness, we believe the pressure to conduct an experiment added rigor to our design process. For example, we had to become more specific about the learning outcomes we were trying to achieve and had to choose an assessment that would be able to show those outcomes. Further, the effectiveness research helped us to be realistic about what we could accomplish: we decided it was unreasonable to deploy CSCL into school with severe behavioral problems and that there would be little benefit for students to collaborate if they had already substantially mastered the target content.

Yet, we did not find that conducting our research in the frame of an experiment limited the contextual observations we could make. Indeed, we collected rich field notes and these led to insights on issues ranging from technical difficulties, to aspects of the Cooperagents materials needing improvement, to the weakness of one particular activity design. In the style of design-based research, we were able to use these findings to rapidly iterate our design even without rigorous statistical confirmation that the underlying factors were causally implicated in the student learning outcomes.

In design-based research, a successful innovation is viewed as a joint product of the designed intervention and the context (Design-Based Research Collective, 2003), and we saw that the combination of both effectiveness and contextual components in our research enabled us to rapidly improve the TechPALS materials and measure a statistical significant improvement in learning. As our work has shown, these components do not need to be in opposition and combining them can be fruitful.
Overall, although we applied the ILD framework in retrospect, we found it was useful in organizing our thinking about how design research can lead to effective CSCL practices. We find it particularly helpful that the ILD framework organizes a trajectory from early stage research to evaluation of broader impacts. Given the complexity of schools, designers must rapidly accumulate many types of insights in order to achieve the goal of making an impact. However, merely accumulating insights and describing improvements to designs feels unsatisfactory as a measure of impact without an explicit comparative measurement step. Experiments provide this comparative measurement of impact and their integration in the ILD framework thus provides a way to keep designers accountable not only to insights and enhancements but also towards progress towards broader, measurable improvements.

Acknowledgements

An earlier version of this paper received the best conference paper award at the CSCL 2009 conference (Rhodes, Greece, June 2009). The International Society for the Learning Sciences granted permission to publish this article.

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305K060011 to SRI International. The opinions expressed are those of the authors and do not represent views of the U.S. Department of Education.

We thank Hewlett-Packard for providing vital support, through its Global Philanthropy program, with a generous donation of more than 100 iPAQ Pocket PCs. We thank our development partners at Eduinnova; their work on the TechPALS project was supported by a grant (CONICYT-FONDEF D04T2036 & CONICYT-FONDECYT 1080100) from the Chilean government. We are also grateful to the teachers, students, and school leaders who participated in this project.

References