Instr Sci (2018) 46:11–33
https://doi.org/10.1007/s11251-017-9433-1

Searching for buried treasure: uncovering discovery in discovery-based learning

Kiera Chase[1] · Dor Abrahamson[1]

Received: 13 March 2017 / Accepted: 27 October 2017 / Published online: 9 November 2017

Abstract Forty 4th and 9th grade students participated individually in tutorial interviews centered on a problem-solving activity designed for learning basic algebra mechanics through diagrammatic modeling of an engaging narrative about a buccaneering giant burying and unearthing her treasure on a desert island. Participants were randomly assigned to experimental (Discovery) and control (No-Discovery) conditions. Mixed-method analyses revealed greater learning gains for Discovery participants. Elaborating on a heuristic activity architecture for technology-based guided-discovery learning (Chase and Abrahamson 2015), we reveal a network of interrelated inferential constraints that learners iteratively calibrate as they each refine and reflect on their evolving models. We track the emergence of these constraints by analyzing annotated transcriptions of two case-study student sessions and argue for their constituting role in conceptual development.

Keywords Design-based research · Discovery learning · Early algebra · Technology

Telling a kid a secret he can find out himself is not only bad teaching, it is a crime. Have you ever observed how keen six year olds are to discover and reinvent things and how you can disappoint them if you betray some secret too early? Twelve year olds are different; they got used to imposed solutions, they ask for solutions without trying.

(Freudenthal 1971, p. 424)

✉ Dor Abrahamson
dor@berkeley.edu

1 Graduate School of Education, University of California, Berkeley, 4649 Tolman Hall, Berkeley, CA 94720-1670, USA

Background and objectives

Discovery-based pedagogical design instantiates Piaget's theory of learning

Jean Piaget believed that "children are active thinkers, constantly trying to construct more advanced understandings of the world" (1968, p. 45). Piaget's conviction, derived from a career of clinical experiments, epitomizes his theory of genetic epistemology, a theory of learning that in turn has inspired an educational approach known as constructivism. Within the domain of mathematics education, constructivists believe that students learn best when conditions enable them to arrive via guided problem solving at normative cultural–historical notions and procedures. Indeed, constructivist educational scholars have contended that teaching young children arithmetic algorithms directly is harmful—it denies them a chance to ground the content each in their own subjective understandings while, even worse, training them to ignore their own thinking (Kamii and Dominick 1998). A protégé of Piaget, Seymour Papert asserted that students best construct concepts when they construct artifacts, whether concrete or virtual (Papert 1980).

Since Piaget, some steps have been taken towards implementing constructivist pedagogy in Science, Technology, Engineering, and Mathematics (STEM) classrooms. This trend is led by scholars, designers, and teachers seeking alternatives to the standard instructional sequence wherein the teacher demonstrates a set of procedures and then the students practice these procedures ("tell and practice," see Schwartz et al. 2011). Airing his concerns with this approach, Catrambone (1998) notes, "Students tend to memorize the details of how equations are filled out rather than learning the deeper, conceptual knowledge" (p. 356).

Alternatives to "tell and practice" are designed for students to solve problems by way of first inventing procedures (Roll et al. 2011), discovering critical features of the problem (Kamii and DeClark 1985), formulating incipient models of the situation (Gravemeijer 1999), or analyzing their peers' work (Kapur 2010), and only then receiving explicit instruction, whether from human instructors or technology-enabled automated supports (Chase and Abrahamson 2015). These discovery-based approaches each uniquely attempt to foster experiences wherein the student actively constructs his or her understanding of a situation through task-oriented interactions with pedagogical resources carefully selected or created for these activities. Specifically, the activities are typically crafted so as to facilitate students' reinvention of essential STEM principles by way of engaging their naturalistic inquisitiveness; soliciting, challenging, and surfacing their implicit assumptions and situated know-how; and then offering and negotiating with them new forms of organizing and representing their actions and reasoning.

Discovery-based learning is, by definition, the process of mobilizing one's agency in becoming aware of a new idea or developing a new skill (Jong and Joolingen 1998), such as when engaging successfully in the types of activities described above. Catering for this process, the educator must bear in mind not only a particular educational point of arrival but also the students' point of departure and ports of call along the way. Just how this is accomplished by educators as well as the efficacy of this approach for students is still a matter of debate in the STEM education research literature (Kirschner et al. 2006; Klahr and Nigam 2004). This paper offers a view on mathematics discovery learning, a view we have found promising in light of its empirical corroboration of pedagogical consequences. From this view, we propose to uncover what we mean when we talk about discovery in discovery-based learning.

Developing conceptual transparency for early algebra: a design research project

The thesis of this paper is situated in a larger research program investigating issues of design, teaching, and learning in mathematics education. In the particular research project motivating the thesis presented herein, we have been seeking to evaluate a general design architecture for discovery-based learning. At the center of the project is an activity design for basic algebra mechanics. The design is called Giant Steps for Algebra (Chase and Abrahamson 2015). Initially, the Giant Steps for Algebra activity design was implemented in the form of concrete objects (Chase and Abrahamson 2013). Later, it was re-implemented in a dedicated microworld (Chase and Abrahamson 2015).

Within the Giant Steps for Algebra (GS4A) activity design, participants are presented with a narrative about hidden treasure and are tasked to find the treasure through modeling the narrative, using available interface functionalities, in the form of a depictive diagram, which they can adjust (see Fig. 1). Their diagrammatic solution can be viewed as a situated proto-algebraic proposition. Note, in Fig. 1, that each giant step subtends exactly 3 "meters." Students infer that the treasure is buried 11 meters away from the Start flag.

In analyzing students' interactions with the instructional materials available in GS4A, we sought to understand how particular materials led to students' apparent understandings of early algebra principles. Similar to Martin and Schwartz (2005), we theorized the students' manipulation actions as cases of adapting the environment so as to extend and distribute the prolem-solving process over the available media (see also Kirsh 2010). We sought thus to articulate students' alleged implicit know-how that becomes expressed and elaborated through these actions (cf. Karmiloff–Smith and Inhelder 1975 on "theory-in-action"; Ryle 1945; Vergnaud 2009 on "theorems-in-action"; see also the philosophical work of Brandom, in Bakker and Derry 2011).

As we explain below, our cycles of empirical research efforts to understand children's cognitive development of mathematical concepts led us to attempt to delineate this type of modeling-based know-how fragments; fragments that, once coordinated, appear to capture a major piece of what we would be comfortable as calling a know-how-based conceptual system undergirding some particular mathematical concept, such as algebra.

"A giant has stolen the elves' treasure. Help the elves find their treasure! Here is what we know. On the first day, a giant walked 3 steps and then another 2 meters, where she buried treasure. On the next day, she began at the same point and wanted to bury more treasure in exactly the same place, but she was not sure where that place was. She walked 4 steps and then, feeling she'd gone too far, she walked back one meter. Yes! She found the treasure!"

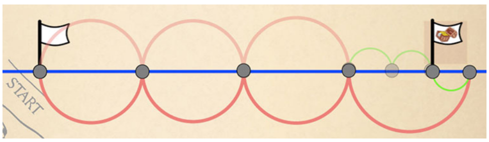


Fig. 1 A sample narrative in the Giant Steps for Algebra (on the left), and a student's representation of this narrative using the GS4A modeling platform (on the right). This narrative instantiates the algebraic proposition of $3x + 2 = 4x - 1$. On both Day 1 (above the line) and Day 2 (below the line) the giant traveled from the Start flag on the left to the End flag on the right. Red arcs represent giant steps (the variable x), green arcs represent meters (the integer units). Students use interface tools to build and revise the diagram

In theorizing students' conceptual understanding of mathematical subject matter as grounded in their modeling know-how, we are suggesting a view of mathematical knowledge as a knowledge of how mathematical procedures "work," just as some children know how bridges, stop-action movies, or computer programs work *because they built those artifacts themselves*. This view of mathematical procedures as artifacts a student can build and use is inspired by Goldstein and Papert (1977), who called to facilitate student learning by creating conditions wherein concepts are "glass boxes" rather than "black boxes," thereby enabling the learner to build subjective transparency.

Transparency is a theoretical construct premised on a sociocultural epistemology, by which conceptual knowledge is deeply implicated in the pragmatic actions that agents perform using relevant cultural tools. As an analytic device, the construct of transparency enables educational researchers to characterize a student's emerging content learning by tracking her evolving understanding for how specific elements of the mediating artifacts she is generating and using, such as mathematical representations, procedures, and manipulatives, function to facilitate the accomplishment of task goals (Hancock 1995; Meira 1998). The construct was developed by scholars inspired by constructivist perspectives. In particular, the construct has been utilized by educational researchers to support arguments for the effectiveness of pedagogical designs that enable students to build and tinker with artifacts. The construct therefore appeared to be an appropriate choice for our own project, as we moved forward.

In the course of analyzing empirical data of students working in this environment, we applied the transparency construct to develop a principled analytic system for implicating relations between students' actions with artifacts and their understanding of the content. This effort led to the articulation of a set of situated micro-skills—bits of context-specific know-how—that the students apparently elicited, developed, and refined as they utilized available artifacts to solve assigned problems. That is, we were implicating students' conceptual transparency of early algebra mechanics by determining their emergent hands-on routines for using elements of the microworld as means of accomplishing localized task goals (see also Gravemeijer 1999). We conjectured that this set of situated know-how captures much of what we mean when we say a student understands a concept. We called each routine a *situated intermediary learning objective*, or *SILO* for short.

Our choice of the term 'situated' is in dialogue with the theory of situated cognition (Greeno 1998), in the sense that we acknowledge the inherent situated quality of human reasoning and learning, as well as with mathematics education research that has availed of situated-ness to support problem-solving specifically for Algebra (Walkington et al. 2013). For sure, Piaget's genetic epistemology corroborates the notion of situativity by virtue of insisting on the emergence of knowledge from goal-oriented interactions in the natural *and cultural* ecology (e.g., his work on moral development through engaging in playground games, Piaget 1965). As educational designers, we conceptualize the engineering of learning environments as creating theoretically informed and developmentally appropriate opportunities for the emergence of situated knowing in line with our pedagogical objectives. As such, learning environments can be perceived as just that—ad hoc micro-ecologies set up by cultural agents explicitly so as to simulate opportunities for students to construct new understanding through exploration and adaptation. As von Glasersfeld (1992) insisted, teachers play important roles in mediating individual students' adaptive construction of knowledge. More generally, we are urged by Cole and Wertsch (1996) to transcend any would-be incompatibility of Piagetian and Vygotskian perspectives so as to formulate a more complete understanding of mathematical ontogenesis in the sociocultural context. For further discussion that problematizes naïve interpretations of situated-ness as they pertain to educational praxis, see Lave (1992) and Brousseau (1997).

Situated Intermediary Learning Objectives (SILOs) in Giant Steps for Algebra

The *SILO—Situated Intermediary Learning Objective*—is a hypothetical construct that characterizes the nature, process, and pedagogical objectives of students' emerging content understandings within discovery-based activities. SILOs thus offer designers, teachers, and analysts traction on the phenomenon of students' discovery-based learning. SILOs are intermediary in the sense that they are very much bound to the environment whence they are enacted, even as they are sufficiently schematic so as potentially to enable reflection, reification, and transfer. In that sense, they are much like situated abstractions (Noss and Hoyle 1996). Notably, when we label students' know-how, we do not imply that the students would describe their knowledge as such. Rather, the SILOs function for the design-research analyst as a means of perceiving in students' pragmatic operations with the available artifacts the emergence of the activity's target content knowledge. Reciprocally, our domain analysis (of algebra) serves to offer candidate "things students should know" about the domain, and we search in our data for possible signs of these things as marking the development of SILOs. For researchers investigating a new content domain through implementing a new design, the set of SILOs coalesces iteratively through qualitative analysis. In turn, the set of SILOs serves the researchers as filters for sifting through the empirical data to foreground and monitor a web of ontogenetic veins leading from the students' first encounter with the task and through to mastery (cf. Pirie and Kieren 1994). Students act on features in their environment (the GS4A interface) and then interpret their actions, a process of physically distributing their thinking (Martin and Schwartz 2005) through epistemic action (Kirsh 1996) and then responding to these enacted 'theories-in-action' ("see–move–see," Bamberger and Schön 1983, 1991). Observing this process enables us to examine students' iterative micro-experiments that gradually contribute to the emergence of the set of SILOs that constitute their knowledge state with respect to the design's target content.

Qualitative analyses of students' behaviors, including their actions and multimodal utterance during the activity as well as in post-intervention assessments, led us to articulate three SILOs in GS4A:

1. *Consistent measures* All variable units (giant steps) and all fixed units (meters) are respectively uniform in size both within and between expressions (days);
2. *Equivalent expressions* The two expressions (Day 1 and Day 2) are of identical magnitude—they share the "Start" and the "End" points, so that they subtend precisely the same linear extent (even though the total distance traveled may differ between days, such as when a giant oversteps and then goes back);
3. *Shared frame of reference* The variable quantity (giant steps) can be described in terms of the unit quantity (meters).

Note how we have attempted to articulate the SILOs in language that is both sufficiently situated so as to index specific operations within a particular microworld and sufficiently general so as to offer potential insight into analogous work on other early algebra problems. As cognitive constructs, the SILOs thus hover in an epistemic space between the locally pragmatic and the conceptually generative. In articulating the set of SILOs that emerge from a single activity design for some target content domain, we cannot claim with conviction that these particular SILOs encompass everything one needs to know about that domain, yet these SILOs do appear to capture concisely much of what students need to figure out within our particular environment: they distil the GS4A problem-solving mechanics that constitute the student's subjective transparency of the target concept, *even as the students are literally and mentally constructing the concept*.

The above synopsis of our earlier work brings us to the present. Until now, we have conceptualized the SILOs as a set of independently emerging situated micro-routines for modeling a problem narrative. In the current study, we explore the possibility that the SILOs are not only a set but a network. That is, we now believe that though these things-to-know can each be described as logically independent, their subjective construction is ontogenetically interdependent. In particular, we will argue that the SILOs emerge iteratively, with each SILO constraining the emergence and calibration of the other SILOs. We are thus revisiting our empirical data with an eye on monitoring the gradual networking of a coherent, inter-calibrated SILO knowledge state.

The Methods section, below, will offer more context on the study and then further explain the idea of an emerging network of constraints. Then in the Results section we will present qualitative analyses of two juxtaposed case studies of participant students working under different experimental conditions. We will present these empirical data in the form of a chain of screen-capture images to depict students' diagrammatic actions as mediated by the hypothesized state of their SILO network. The Discussion section will summarize our findings to argue for a view of discovery-based mathematics knowledge as a network of mutually constraining situated know-how emerging from iterated actions of problem-solving construction and reflection. Finally, the Conclusions section will offer closing comments and implications for educational theory and practice.

Method

This paper reports on results from a follow-up re-analysis of a corpus of data generated in the context of a design-based research project. Due to space constraints of this journal issue, we refer readers to our earlier publications for information about the GS4A design rationale, where we cite the educational-research literature informing our work. Those publications also detail our design-based research method, including interview and analysis methodology. The current paper will present only as much details as is necessary for the reader to contextualize the episodes we report and analyze in this particular study.

The original project was conducted in the design-based approach to empirical research. This approach combines a framework for engineering products through iterative design cycles with methods for inferring generalizations germane to the science of learning (Cobb et al. 2003; Edelson 2002). Working in this iterative process enabled us to investigate the domain of algebra in search of potential explanations for why it is difficult for students to access, ultimately developing a conjecture that led to a proposed educational design, Giant Steps for Algebra. The empirical data discussed herein were collected using a task-based semi-structured interview protocol (Clement 2000; Ginsburg 1997). These data were treated using micro-analysis techniques (diSessa 2007; Kuhn 1995; Parnafes and diSessa 2013; Siegler and Crowley 1991).

A total of 40 Grade 4 and Grade 9 students participated individually in the study. Within each age group, participants were randomly assigned to one of two conditions, Discovery and No-Discovery, balancing as much as possible for gender as well as for ability levels as reported by their teachers. In both conditions students were asked to solve the same problems using an open number-line-like representation that is grounded in the work of Dickinson and Eade (2004). For the Discovery students, modeling was labor intensive, because the computer did not scaffold the production of any of the SILOs. Only after the students had articulated (either via verbal or repeated production) how a particular SILO operates did the interface 'level-up' and automate this functionality for the students. For example, Discovery students had to toil to make the giant steps equivalent (SILO 1), but once they phrased that principle, the tutor took measures so that the computer enabled them to adjust the size of giant steps uniformly. These design choices are the operationalization of our pedagogical framework for building subjective transparency. We refer to this pedagogical approach as Reverse Scaffolding (Chase and Abrahamson 2015). No-Discovery students, on the other hand, received ab initio the automatic implementation of all the SILOs (see Appendix A for more details about each level).

After completing the set of 9 GS4A problems (see Appendix A), all participants solved a set of 5 post-intervention assessment items (see Appendix B). These items were designed to measure the participants' subjective transparency for the structural properties of the problem, that is, to measure their attainment of, and fluency with the SILOs identified through this project. The items consisted of New-Context problems designed to measure for the application of learned skills (transfer) and In-Context problems that targeted the SILOs directly within the familiar GS4A setting.

Once all the data were gathered, we employed interaction-analysis techniques (Jordan and Henderson 1995) to search for patterns across the data corpus. Using the video and screencast data, we were able to code participants' interactions with the GS4A interface in terms of their knowledge per the set of SILOs (see below). The SILOs also served as scoring criteria for measuring individual students' achievement on the post-activity assessments, with each SILO further graded into achievement levels.

For the current study, we hoped in particular to implicate students' micro-behaviors as driven by their insights into properties of the model they were building. Drawing on a study of architects' modeling process (Schön 1992), we expected to observe chains of "see–move–see." That is, you observe what you have created, make a change based on your observation, and then observe the results of the change, and so on. We hoped thus to characterize each change in students' knowledge state as evolving through inferred logical micro-implications from perceived structural features in the virtual diagrams they themselves were creating. As such, the notion of knowledge state is akin to a family of theories, such as theory-in-action (Karmiloff-Smith and Inhelder 1975; Vergnaud 1994, 2009) and cognitive anthropological analyses of problem-solving with material models (Bamberger and Schön 1983). The term 'state' is a theoretical construct that should connote a current configuration of a multi-node knowledge structure—a configuration that changes iteratively in response to students' new insights. We conjectured that these diagrammatic constraints on modeling action would operate not only within the SILOs but across SILOs, so that we could argue for conceptual growth as the dynamical and systemic coalescing of a network of mutually constraining modeling micro-routines. By tracking patterns of inter-constraint calibration within each study participant, we would then compare these patterns across participants in each condition.

We developed a coding system to represent micro-events of students calibrating the SILOs (see Fig. 2). We encode each shift in a participant's diagrammatic goal as an *event*. Within an event, each SILO is represented by a circle. In this study, there are three SILOs. SILO 1 is an exception—it is subdivided into two parts during the activity, as corresponding to the diagrammatic units in question, meters and giant steps. Each SILO can be

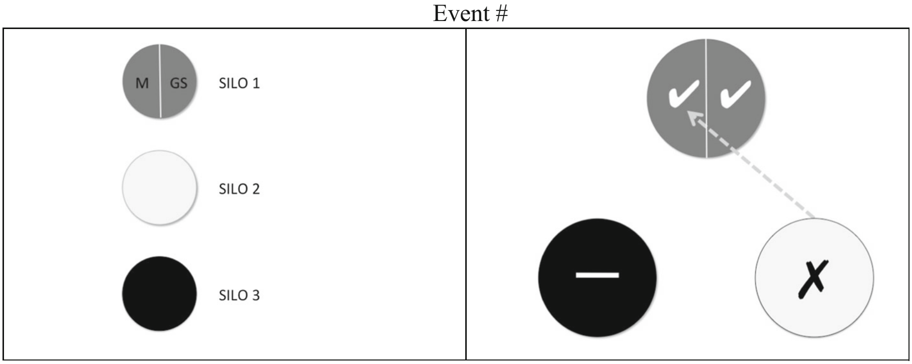


Fig. 2 Example of the analytic representation format used in the analysis section. On the left: a legend of three SILOs. On the right: a particular configuration of these same three SILOs. Each SILO is either inactive (—), implemented (✔), or violated (✗). The arrow depicts an emergent disequlibrium informing a diagrammic goal within the GS4A interface, which the student then acts upon (violation of SILO 2 implicated violation of SILO 1 [meters], which the student then corrected by adjusting the location of screen elements, all the while operating within the constraints of SILO 2)

marked so as to indicate whether students' behaviors have been analyzed as contravening that SILO (✗), adhering to that SILO (✔), or ignoring that SILO (—). Thus each event is a snapshot of the participant's implicit working hypotheses at that moment, their knowledge state. An arrow indicates when a participant's action expresses a particular SILO yet is constrained by another SILO, as if characteristics of one SILO are influencing another SILO, where the arrow signifies the directionality of this influence. For example, the image in Fig. 2 encodes some event where, to our analysis, the participant has looked at her diagram and is alerted to a logical or practical problem relevant to SILO 2 (she detected in her diagram unequal spatial extents for Day 1 and Day 2). To correct this problem, she reenacts SILO 1 (e.g., she uniformly adjusts the spatial extent of all the "meters" in the diagram) as constrained by SILO 2 (ensuring equivalent total extents for the two "days"). We observed patterns in the network's evolution, where most often SILO 2 influenced a re-enactment of SILO 1 as described above. The arrows describe the participants' emerging implicative association between SILOs. The network consolidates through recurrence, evaluation, and articulation of these implicative associations.

By encoding the videography of individual students' work in the GS4A learning environment as a sequence of events, we intended to evaluate whether the modeling process is a concatenation of see–move–see iterations. Such a finding would enable us tentatively to expand our claims about how discovery learning transpires, at least in this activity design.

Results

As reported in earlier publications (Chase and Abrahamson 2015), quantitative analysis revealed the main effect that participants in the Discovery condition achieved better than their No-Discovery peers on the post-activity assessment items. These results were persistent across all three post-assessment categories (In-Context, Transfer, and Overall). We concluded that the discovery-based activity architecture increased student outcomes in our study. Qualitative analyses suggested that Discovery students' conceptual advantages could be attributed with specificity to their discovery experiences: what they built, they understood (Chase and Abrahamson 2015).

The remainder of this section presents results from our new round of qualitative re-analysis on the data corpus. We present two event-based comparisons of students' modeling process in the Discovery and No-Discovery conditions. The vignettes are organized as matched pairs, with compatible Discovery and No-Discovery participants who were working on the same items. This juxtaposition will serve to highlight critical differences relevant to each study condition. For the comparison, we have selected two 4^{th} grade participants who each exemplifies behaviors typical of their respective study condition and ability group. Both participants were characterized by their teacher as performing at mid-level in mathematics.

Item 2: comparison

We will begin with Frankie, who is working in the Discovery condition. She is working on Item 2, where the narrative instantiates the algebraic proposition $3x-3 = 2x + 2$. We will begin by unpacking each Event in the form of a more detailed annotated transcription, so as to warrant our inferences regarding Frankie's succession of knowledge states.

Like many of her study-condition peers, Frankie began by building giant steps and meters of approximately the same size and relationship as in the previous item (see Fig. 3). Having almost completed her first attempt, Frankie realized that her model was problematic. Event 1 captures her realization of dissonance between her current approach to solving the problem and some unintended consequences. In particular Frankie realizes that by pursuing her current plan of action she would impugn SILO 2. Event 1 signifies a new knowledge state that will result in Frankie generating a new diagrammatic goal moving forward.

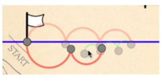
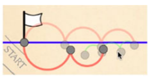
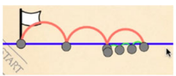
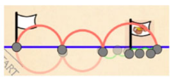
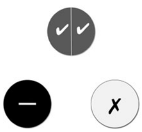
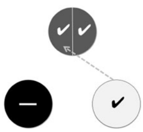
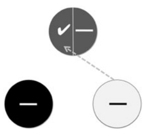
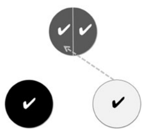
Frankie Item 2				
Event #	1	2	3	4
Model				
Utterance	Frankie: I'm already past the end [her cursor scrolls over the end of the Day 1 model, which her Day 2 model has already skipped over].	Frankie: Yeah, so the meters are probably…. [her cursor scrolls from the end node on Day 1 towards the right, and circles where the end of Day 2 could be].	Frankie: Can I go back to day 1? Res.: Yeah, you can do whatever you want. Frankie: [she begins by deleting Day 2 of her model, then returning to Day 1, and remodeling the meters much smaller]. Res.: What are you trying to do? Frankie: I'm trying to make it look…I don't know Res.: It looks like you're trying to make the meters smaller. Frankie: Yeah.	Frankie: So how many meters did she bury it from the starting point? I think this time it is 4. Res: Really? Frankie: Oh no, 5, 5 meters.
Knowledge State				

Fig. 3 Complete overview of all Events for Frankie during Item 2

Frankie's model in Event 1 accurately represents giant steps and meters each as a consistent measure, therefore SILO 1 is encoded with a checkmark. Her utterance and the actions taken with the cursor indicate her realization that SILO 2 could thus not be achieved; therefore SILO 2 is encoded with an X. Frankie does not attend to SILO 3, and so it is encoded with a dash. Looking at Frankie's diagram reveals that during Event 1 she modeled a spatial relationship where 1 giant step is approximately 2 meters. This implicit hypothesis is surfaced and called into question once Frankie imagines how her model would look if completed. Frankie must react to this anticipated error by generating a revised diagrammatic goal. Thus, per our analytic conjecture, perceptual features in a student's diagram may iteratively constrain the evolution of another SILO. Immediately following, Frankie makes a small action indicating that she is considering a possible repair to the violation that had just occurred. This is captured in Event 2. She moves her cursor from where the current model ends on Day 1 to a location where Day 2 could end, thus tentatively reconciling SILO 2. This is encoded as an arrow indicating the new constraint that satisfying SILO 2 places on Frankie's diagrammatic goals, specifically how she models the meters.

Frankie then begins making this imaginary projection into reality (see Kirsh 2009), as captured in Event 3. First, she must remove some elements from her current model, including all of Day 2 material. As for Day 1 material, she keeps the giant steps intact but alters the meters.

Frankie's revised hypothesis is calibrated to the new constraint she has just discovered. Her new diagrammatic approach is consistent with her previous approach, with only one slight adjustment: She must modify the spatial relationship between the meters and the giant steps. Frankie continues to model her meters consistently. But in order to satisfy SILO 2, she keeps making her meters smaller and smaller, toggling between Day 1 and Day 2, until the days' respective ends are finally collocated (see Event 4). Frankie then correctly calculates the location of the buried treasure. In this vignette we see that Frankie vacillates between attending to the size of the meters (SILO 1) and the relative spatial extent of the two days (SILO 2). Frankie does not change or comment on the size of the variable (SILO 1: giant steps)—this diagrammatic figure appears to be an immutable feature of her model. Rather, we have witnessed how a breakdown in implementing the diagrammatic constraints of equivalent expressions (SILO 2) stimulated Frankie to re-implement and thus calibrate the principle of consistent measures (SILO 1—meters).

As a point of comparison, Zula is working in the No-Discovery condition (see Fig. 4). Bear in mind that in this condition the computer is automating the production of consistent giant steps and meters. At first glance, it may appear as though there are many similarities between Zula and Frankie, however, the differences between the students become evident as the activity progresses. Zula is also working on Item 2 ($3x - 3 = 2x + 2$). She models the entire narrative. She then determines that her model cannot be accurate because the Day 1 and Day 2 end nodes are not collocated. This is a similar realization to that of Frankie.

In order to repair the impugned SILO 2, Zula contracts the Day 2 model to the left so as to near its end node to the Day 1 end node. However, this action does not result in any new information, as the model still does not satisfy the SILO 2 requirement.

Then Zula pauses and expands the size of the giant steps, which auto-expands the entire model, until the two days' respective ends are collocated. Event 3 introduces one constraint, as Zula must adjust her diagram to satisfy SILO 2. She can accomplish this diagrammatic goal by changing the size of the giant steps. This is encoded as an arrow from SILO 2 to SILO 1 in Event 3. Once Zula has repaired her model, she is able to

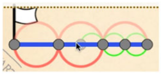
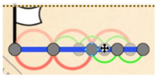
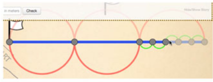
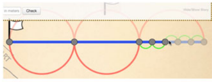
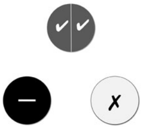
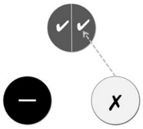
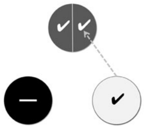
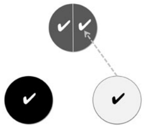
Zula Item 2				
Event #	**1**	**2**	**3**	**4**
Model				
Utterance	Zula: That wouldn't work. Because she buried the treasure here [scrolls over the final node on Day 1].		Zula: You need to make it like that [pauses]. No, like there [stretching] there [stretching], there.	Zula: On Day 1 she goes back 3, and on Day 2 she goes forward 2, and they have to line up…. [Reads] "so how many meters away is the treasure buried from the starting point"…. If there were 5 [scrolls over the giant step that is subtended by 5 meters], then 5, 10, 15. [She enters 15 and gets feedback that this is incorrect]
Diagram				

Fig. 4 Overview of all Events for Zula during Item 2

calculate the size of the giant steps and determine a correct solution. Event 4 captures this process as well as her struggle to calculate an accurate solution.

After entering "15" as a solution, Zula returns to her model and eventually calculates 12 as her solution. Zula had neglected to account for the negative integer on Day 1, as the giant went *back* 3 meters. In this vignette we saw that Zula only has one occasion where her actions result in feedback that is contrary to her anticipated outcomes. This occurred in Event 2. We may tentatively submit that the No-Discovery condition may offer reduced occasion for participants to engage in a see-move-see sequence that results in inter-calibration of the SILOs.

Item 4: comparison

In this next section we will compare the work of Frankie and Zula on Item 4, a narrative that corresponds with the equation "$3x + 2 = 4x - 1$". We will begin with Frankie. She has moved into the Level 2 of the Discovery condition (see Appendix A for details about the levels). In Level 2 the interface generates equivalent meters automatically but the user still controls the size of each individual giant step. Frankie begins by modeling Day 1 entirely and then advances to Day 2 and models only the giant steps (see Fig. 5). In this initial pass at solving the problem, Frankie is focusing on generating consistent giant steps across Day 1 and Day 2 of her model.

Before she even completed the model Frankie realized that if she continued to build this model using her existing parameters, once completed it would violate SILO 2, "equivalent expressions." This infraction propels a revised diagrammatic goal.

Frankie's next move, captured here in Event 2, is to delete what she had modeled for Day 2 and begin again. Her reconstructed model no longer impugns SILO 2. This is encoded with the arrow in Event 3, Fig. 5. However, her revised approach involves changing the giant steps on Day 2, so that they extend further to the right (see Fig. 5). Her attempt to repair Event 2 resulted in a new violation. Frankie's revision now fractures her

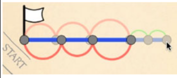
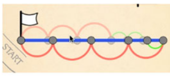
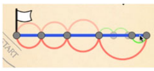
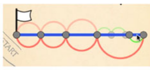
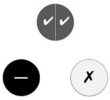
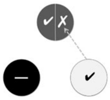
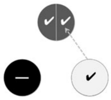
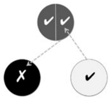
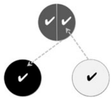
Frankie Item 4					
Event #	1	2	3	4	5
Model					
Utterance	Frankie: Well she is in the same spot. [Frankie moves her cursor in an arc, where the 4th giant step could go. Then she hovers her cursor over the last node in her Day 1 model, indicating that if she took a 4th giant step forward, as the story suggests, and 1 meter back, the model would not work].	Frankie: Ok, so that is about the same [referring to the collocation of the Day 1 and Day 2 ends]. So how many meters? [She begins to calculate the total distance from the start in meters].	Frankie: Well how do I make them bigger? [she hoovers over the giant step on the far right, because this one is much bigger than the others.]	Frankie: These look like 3 meters, and these look like 2 meters. Res: Did the story say anything about extra small giant steps? Frankie: No, I just drew them like that. So actually 3 meters	Frankie: No, I just drew them like that. So actually 3 meters. [She then calculates the correct response]
Diagram					

Fig. 5 Overview of all Events for Frankie during Item 4

notion of consistent measures. It takes Frankie some time to recognize why this model (Event 3) is faulty.

After some unsuccessful attempts to calculate the distance in meters, Frankie sees that her model cannot work and begins moving the giant-step nodes on top of each other. Event 3 returns Frankie to a newly calibrated inter-constrained execution in which SILO 1 can be corrected without disrupting SILO 2.

Then Frankie attempts to calculate the size of a giant step in meters. In Fig. 5, Event 4 we see that Frankie attempts to establish a shared frame of reference, but that her diagrammatic imperfections briefly cause some confusion. Almost immediately following, in Event 5, we observe that she sees beyond what she produced to her intended diagram. She establishes the shared frame of reference, SILO 3, and is able to calculate the correct solution.

We now join Zula working on Item 4 (see Fig. 6). She begins by modeling the complete narrative and then looks at it and determines that it is not accurate. This change in her knowledge state is captured in Event 1. Zula then stretches the model until the ends are collocated, as seen in Event 2. She then focuses on the shared frame of reference (SILO 3), establishes the size of a giant step, and calculates a solution to the problem, as seen in Event 3.

Zula is not confronted with any feedback from her model or her modeling actions that would implicate a particular violation or necessitate a diagrammatic repair. While she does not immediately satisfy the criterion for establishing equivalent expressions (SILO 2), she can do so without experiencing further constraint on or from other parts of the network.

Post-activity assessment item comparison

Up to this point in the comparison of Frankie (Discovery condition) and Zula (No-Discovery condition), we see that there are discernable differences; in particular Frankie's

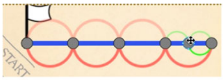
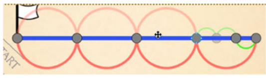
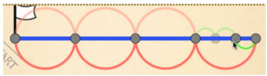
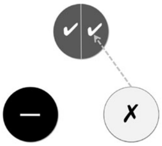
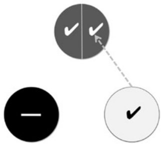
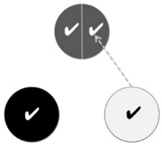
Zula Item 4			
Event #	1	2	3
Model			
Utterance	Zula: She did not bury it in the same place.	Zula: So, have to fix it [she stretches the model until the ends are collocated]. There.	Zula: So 3 [points to the final giant step that is subtended by 3 meters] 6, 9. No wait, she buried her treasure here [points to the shared final node] so 2, plus 3 is 5, plus 3 is 8, plus 3 is 11. 11.
Diagram			

Fig. 6 Overview of all Events for Zula during Item 4

event sequences are longer and more varied. It is during the post-activity assessments that these differences become more evident. Frankie is working on one of the post-activity assessment items, the Two Buildings problem (see Appendix B). She begins by reading through the problem and then points to the text about Building A that has 10 floors and a 20 foot spire. (Building B, which is equally tall, has 11 floors and a 10 foot spire.)

Frankie's initial attempt at establishing equivalency between the two buildings is to vary the size of what should be conceptualized as a consistent measure (see Fig. 7). Thus, Event 1 captures her initial knowledge state.

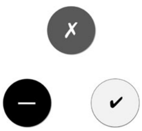
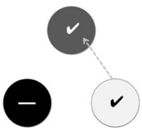
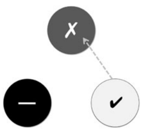
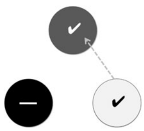
Frankie Two Buildings				
Event #	1	2	3	4
Model				
Utterance	Frankie: These floors [building A] will have to be bigger than these floors [points to the text describing building B]. Res.: Why? Frankie: Because 11 and 10. Wait, Floors are like stories? Res.: Exactly, like when you go up in an elevator… Frankie: So they don't have to be the same size.	Frankie: Actually they do [referring to the height of the floors]. Frankie: So the building is the same height, or the Spire. Are they talking about the building *and* the spire? Like together Res.: Yeah	Res.: So you're saying that the floors have to be different sizes. Frankie: yeah. Res.: Is that typically how it would be in a building? Frankie: No.	Frankie: [begins rereading the prompt] Well, then, there's, the spire is 10 feet tall. So this one [Building B] would be like half the size of it [Building A]. But this building [referring to Building B] would have to be, let's say this building is bigger [draws another floor] Res.: So it's bigger by one floor? Frankie: Yeah [pointing to the top of Building B with her pencil].
Diagram				

Fig. 7 Frankie's encoding and transcript for the Two Building transfer problem

Immediately afterwards, Frankie revises her approach and decides that the floors do need to be consistent across both buildings. Subsequently, she turns her attention to the overall challenge of establishing equivalency and verifies her interpretation of the problem's components. This is captured in Event 2.

Frankie then begins to diagram her thinking. She starts by drawing Building A. She is unsure how tall to draw the spire. She begins by drawing a second building that is the same height as the first, excluding the spire. In Event 3 we can observe how Frankie reverts to varying the size of the floors between buildings in order to establish overall equivalency.

Immediately after Event 4, Frankie returns to the text. She rereads the narrative and comes to the conclusion that it is possible for the floors to be the same height across each building because the spires are different heights. Frankie redraws the top of Building B (on the right) and adjusts the height of the spire. Frankie has discovered how SILO 1 and SILO 2 can co-obtain.

In Event 5 (see Fig. 8), Frankie looks at each component of her diagram and then determines the shared frame of reference, in this case the height of a floor. This signals a change in Frankie's knowledge state, as exemplified in Event 5. She has determined a shared frame of reference by examining the constraints of her diagram. Immediately following, Event 6, the researcher asks Frankie to explain her thinking. This vignette exemplifies the iterative nature of the inter-constraint calibration that characterizes discovery in this context. The participant often conceptualizes one of the constraints as central, then shifts attention to another constraint and cannot immediately satisfy both constraints. It seems that through the process of recognizing how one of the SILOs has been contravened the participant gains a better understanding of this particular SILO and in particular how it networks the set of considerations for solving this contextual problem. Additionally, Frankie uses her model as a way to check her reasoning and provide herself with feedback that enables her to make the necessary micro-adjustments along the way. This is very different from the behavior we see from participants in the No-Discovery condition, such as Zula.

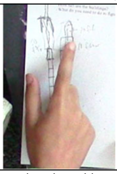
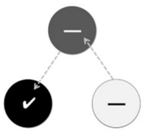
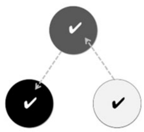
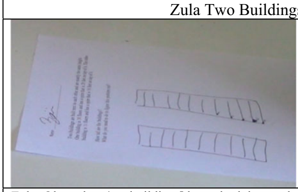
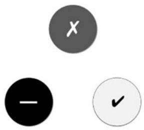
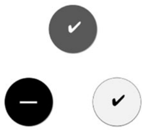
Frankie Two Buildings Event 5	Frankie Two Buildings Event 6
 	
Res.: So you just drew this one bigger by one floor, why? Frankie: Cause it has 11 floors. Res.: And this spire is 10 feet? Frankie: yeah Res.: And this spire is 20 feet? And it goes all the way to here. Frankie: Yup Res.: And then the rest is all the same? Frankie: Yup…then this would have to be 10 feet [pointing to the 11th floor on building B]	Res.: Why? Frankie: Because it wouldn't be the same [points to the spires of the building] Because this is 10 [her pointer finger to tap on the spire of building B] and this is 10 [uses her middle finger to tap on at the 11th floor of building B], and it is the same [moves her fingers over to the 20 foot spire on Building A, and taps them both together], and then it was just be the same on the rest.

Fig. 8 Frankie's drawing, encoding, and transcript Events 5 and 6 in the Two-Buildings transfer problem

Zula Two Buildings Event 1		Zula Two Buildings Event 2
		
Zula: Ok so there's a building [draws both boxes the same height] and it has 1,2,3….[counts out 11 floors on building B and 10 on building A]		Zula: But then this one is higher because it is 11 floors [draws a line across the top of both of them]

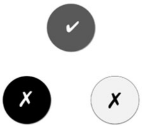
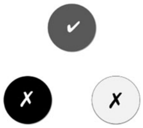
Zula Two Buildings Event 3	Zula Two Buildings Event 4
	
Zula: Then this one [building A] has a spire that is 20. So 20 minus 11 is…... No 20 plus 30, and then 30 minus 11. is 29. No 19. So right now this [building A] is 19 feet taller. And then this [building B] has a spire that is 10. And 10 plus 19 is 29. So they are not the same height.	Res.: How do you know that. Zula: Because if this one is 10 feet [points to building A]…Oh this one is 10 floors. We don't know how big the floors are.

Fig. 9 Zula's diagrams, encoding and transcripts for two buildings transfer problem, all Events

Zula is also working on the Two Building post-activity assessment item (see Fig. 9). She has read through the story and begins drawing out the scenario. Her first move is to draw two rectangles that are the same size, then subdivide them into 'floors' (see Event 1). This diagrammatic goal inadvertently neglects SILO 2.

In Event 2 Zula reminds herself that Building B is indeed taller, because it has 11 floors. She draws a line across the top of both buildings to indicate this. This action brings her focus to SILO 2.

In Event 3 we observe that Zula proceeds by modeling the spires on each building and then performing a series of calculations, leading her to the conclusion that the buildings are actually not the same height, thus negating SILO 2. Her calculations do not attempt to establish a shared frame of reference, thus negating SILO 3.

Immediately following Event 3, the researcher asks Zula to explain her reasoning. This opportunity to interrogate her reasoning helps Zula recognize that she was confusing floors and feet, and this brings her back to calculating the height of each floor (Event 4). Zula returns to her drawing and determines that 19 ft. is the height of each floor, which is incorrect.

In contrast to Frankie, Zula does not use the model she is creating so as to reflect, to calibrate her thinking, or generate a new and more nuanced approach. She does not generate new diagrammatic goals when faced with inconsistencies, which in turn inhibits her development of new knowledge states as the activity progresses.

Discussion

The Giant Steps for Algebra learning environment, including its source materials, assigned task, means of production, and facilitation techniques, was designed with the explicit objective of enabling students to bring to bear common sense in assembling piecemeal what become diagrammatic instantiations of the target content, early algebra, complete with its structural, relational, and functional coherence. That is, we designed for students to build a transparent algebra, where all the procedural operations are logically necessary, given one's tacit know-how about how things are in the world and vis-à-vis the recursively emerging figural constraints of the model they were constructing. Thus, for a moment do we not claim that students discovered algebra "on their own"—claiming that would be a travesty of the entire design philosophy, agenda, rationale, and procedure. On the contrary, we regard the learning process, replete with tutorial interaction, as exemplifying heavy-handed sociocultural mediation of mathematics knowledge. However, this mediation is largely premeditated, embedded in the learning environment. Rather than spoon-feeding forms and dictating procedures, we created conditions for students to encounter cognitive challenges whose pragmatic solution amounted to learning core notions of the target content. Our objective in this paper has been to characterize the piecemeal assemblage of the new conceptual elements, which we call situated intermediary learning objectives (SILOs), not as independent thrusts but as interdependent, iterative, pragmatic adjustments of construction micro-routines vis-à-vis emerging features of the solution models into an inter-calibrated network. The network is composed of two primary features, the nodes and the rays. The three nodes are the SILOs, and the rays signify implicative associations between nodes or how each SILO constrains the other.

The methodological rationale of the study was to hypothesize differences between the learning gains of students who engaged in the experimental Discovery mode of the activity, per our descriptions above, and those who engaged in solving the same problems only that the technological environment relieved them from the efforts of figuring out best micro-modeling practices. As such, No-Discovery students received ready-made tools, in a manner that we believed reflects mainstream-classroom instructional experiences. As researchers, we were intent on evaluating whether our hypothesis would bear out in the form of manifest differences in the learning gains of the two groups (the main effect), and we further aspired to implicate differences in the process that led to these diverging outcomes. Moreover, we hoped that we could learn more about the nature of discovery learning by virtue of grounded-theory characterization of observed differences in process. This paper expressed results of analyzing and juxtaposing case-study accounts of the experimental and control groups' recorded behaviors.

Our quantitative and qualitative analyses supported claims for greater learning among Discovery as compared to No-Discovery students (Chase and Abrahamson 2015). Our further qualitative analyses presented herein now exemplify how in the absence of ready-made interaction functions, Discovery participants were obliged themselves to intuit, infer, determine, construct, *and inter-calibrate* features of the mathematical system. By way of contrast, participants in the No-Discovery condition, who received the interaction features ready-made, did not experience as much insight into the embedded mathematical principles. Indeed, their individual SILOs were not as manifest or articulated and their set of SILOs was not as inter-constraining.

The results of this study further highlight the importance of differentiating between *learning to use a tool* and *using a tool to learn*. In many cases, No-Discovery participants developed a know-how (Ryle 1945) that was situated in the particular immediate context and subsumed the tool. What is more, they could articulate how they were using this available tool to solve the problem at hand: They used the ready-made interface features and often understood the relationship between their mousing actions and changes on the screen. However their understanding was only "screen deep." For example, No-Discovery participants understood that the tool could stretch or shrink giant steps, but not how this action functions within the larger mathematical system. Learning how to operate a tool did not lead to compatible opportunities to develop subjective transparency of the mathematical system. The tool's ready-made utilities were never interrogated with respect to problems of practice that the tools each solved. The utilities were conceptualized as manipulative features, not as solutions. Just as in the case of bicycle gears, one can become highly skilled in using an artifact's utilities without ever questioning their rationale or build, without looking under the hood. Using a ready-made tool does not necessitate the development of situated intermediary learning objectives (SILOs).

By contrast, participants in the Discovery condition had to enact and formulate interaction goals and then compensate for the tool's shortcomings; only once they had articulated these compensatory strategies did the experimenter supplement those repair strategies into the tool as built-in utility features. Through this process, the participants came to understand their own ideas with more clarity and, in so doing, achieve greater fidelity with the target concepts. Thus the Discovery condition enabled participants to encounter a problem of practice, enact their solution, and ultimately articulate and confirm it. The process of first needing a particular tool and only then receiving it created increased opportunities to develop subjective transparency of the emergent mathematical system.

We infer that *learning to use a tool* is a learning activity that does not necessitate that the participants encounter the problem of practice nor articulate how the built-in features of an artifact enable their success. Furthermore, knowing how to operate a tool does not guarantee a seeing and understanding of the cultural–historical disciplinary knowledge embedded in the tool (see also Meira 2002). On the other hand, *using a tool to learn* implies that users build new knowledge by engineering improvements to imperfect tools, where these engineering micro-solutions embody the design's learning objective. In this paper we have exemplified and explicated the process of using a tool to learn. Although the process is premeditated by the designer, for the student this process is not teleological or concept-oriented—it could not be, because the students do not know what the concept in question might be. Rather, the process is problem driven and detail-oriented; with each modeling operation giving rise to material (or virtual) features that, in turn, require modification to modeling routines and emergent heuristics. Conceptual knowledge is a reflective mastery of the sum total of these interdependent heuristics (cf. Bereiter 1985, on the "learning paradox").

This paper contributes to the discovery-based learning landscape in three ways. We propose a pedagogical vision, a heuristic design framework, and an epistemological framework. Firstly, our pedagogical methodology is focused on building subjective transparency as our characterization of discovery. Secondly, we developed a heuristic design framework that outlines a procedure for determining the SILOs underlying a target concept through a combination of top-down cognitive domain analysis and bottom-up micro-ethnographic analyses of clinical interviews. Our activity architecture iteratively implements the target content in a discovery-based learning environment including materials and activities. Finally we have proposed an epistemological framework for characterizing, representing, and monitoring the emergence of mathematical knowledge. These three aspects of the project—pedagogy, design, and epistemology—cohere around the proposed construct of the SILO.

While we used technology to implement this particular design, we believe that it is not necessary. Stepping back, we realize that our use of technology in building the learning environment required of us to articulate a theory of interactive learning, paying close attention to the implications of automation when designing for discovery. However, this theory may well obtain and generalize to other media for mathematics instruction. More generally, this methodical approach to the design and evaluation of instructional environments involving student construction of systems may support research efforts to investigate student STEM learning in activities that explicitly foster and foreground student assembly of functional systems, as in fabrication laboratories associated with the Maker movement.

Limitations

The construct of SILOs emerged in the particular context of this design-based research study. As of yet, we have no evidence to substantiate a claim that other let alone all mathematics content can be conceptualized via its SILOs. Furthermore, even if we could determine a new content domain's SILOs, we cannot as yet state that those specified SILOs would be conducive for inter-calibration, as discussed in this study. Additionally, this study presents an activity that was delivered individually and outside of the classroom, and we recognize that we therefore cannot predict how this activity and inter-calibration would emerge in classroom settings. We sincerely hope that our work will inspire further research and scaling into other domains and settings of mathematics and, perhaps more broadly, STEM education.

Conclusions

In earlier publications, we put forth a framework for educational designers to build interactive technology by which students reinvent mathematical subject matter content through diagrammatic modeling of a problem-based source narrative (Chase and Abrahamson 2015). The framework specifies how to embed within interactive technology a task-oriented activity sequence for the iterative development of conceptual transparency for target mathematical content. This paper extended our framework through qualitative re-analysis of empirical data to characterize students' learning process as iteratively calibrating a set of modeling micro-routines in light of emergent constraints imposed by features of their own diagrammatic models. This network of calibrated constraints both comes to constitute students' understanding for the meaning of the content and organizes their prospective problem solving for appropriate situations. If in previous publications we portrayed students' emerging content knowledge as a set of situated intermediary learning objectives (SILOs) per se, we now can explain how these SILOs inter-coalesce into a tight action-oriented cognitive structure, a network of mutually constraining SILOs. We hope this keener analysis of the learning process will be of use to educational researchers and designers.

These are early days to evaluate the scope and reach of the inter-calibrated SILO model of learning as a means of explaining discovery-based pedagogy, both its theory and practice. To the extent that researchers explore these ideas further, we anticipate the utility of doing so in the context of learning environments that are geared to elicit students' relevant experiential resources, even skills as simple as walking or clapping hands together (Zohar et al. 2017).

Appendix A

Schematic listing of intervention items for both study conditions as well as activity sequence and interface functionality for each condition

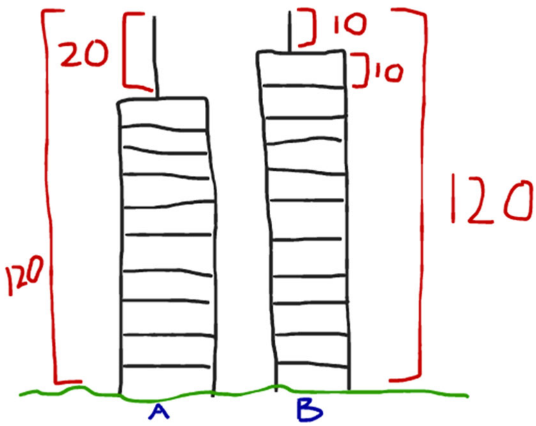
Items		Condition	
Item	Annotated narrative and algebraic expression.	Discovery	No discovery
1	Four giant steps forward. Then three giant steps forward and two meters forward. $4x = 3x + 2$	All manual	All automatic
2	Three giant steps forward and three meters back. Two giant steps forward and then two meters forward $3x{-}3 = 2x + 2$	All manual	All automatic
3	One giant steps forward and then eight meters. Two giant steps and then six meters. $1x + 8 = 2x + 6$	All manual	All automatic
	Transition to level 2		
4	Three giant steps forward and then two more meters. Four giant steps forward and then one meter back. $3x + 2 = 4x - 1$	Manual giant steps automatic meters	All automatic
5	Three giant steps forward and then three meters back. One giant step forward and then one more meter. $3x - 3 = x{+}1$	Manual giant steps automatic meters	All automatic
6	Two meters forward, then two giant steps forward, then three meters forward. One giant steps forward, then one meter forward, then two giant steps forward, then one meter forward. $2 + 2x + 3 = x + 1{+}2x + 1$ $(2x + 5 = 3x + 2)$	Manual giant steps automatic meters	All automatic
	Transition to level 3		
7	Five meters forward and three giants steps back. Two giant steps forward. $5 - 3x = 2x$	All automatic	All automatic
8	Two giant steps forward and four meters back. The one giant steps forward and three meters back. $2x - 4 = x - 3$	All automatic	All automatic
9	Three meters forward, then two giant steps, then four meters. Then two meters and three giants steps. $3 + 2x + 4 = 2 + 3x$	All automatic	All automatic

Appendix B

Post-intervention items for both study conditions

New-context problems

The two buildings problem



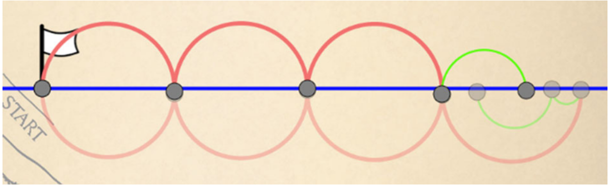
Two buildings are built next to each other and are exactly the same height. One building is 10 floors and has a spire that is 20 feet on top of it. The other building is 11 floors and has a spire that is 10 feet on top of it.

How tall are the buildings?
What do you need to do to solve this problem?

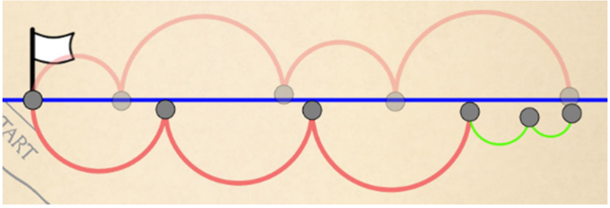
The turtle years problem

My turtle, named Yurtle, is being tricky and won't tell me how old she is.
Help me figure out how old she is in human years.
Yesterday she told me that she has lived 3 turtle years and 2 human years. She also told her 4th turtle year will begin in 3 human years.

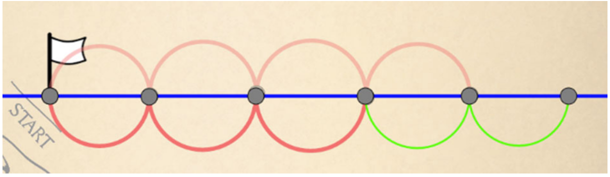
In-context problems



In this screenshot a hypothetical user created inconsistent meters. The study participant must: (a) identify the user's error; (b) redraw the scenario with consistent meters; and (c) determine the treasure's location.



In this screenshot a hypothetical user created inconsistent giant steps (variables). The study participant must: (a) identify the error; (b) redraw the scenario with consistent giant steps; and (c) determine the treasure's location.



In this screenshot a hypothetical user did not represent equivalence. The study participant must: (a) identify this error; (b) redraw the scenario with matching end-points; and (c) determine the treasure's location.

References

Abrahamson, D., Chase, K., Kumar, V., & Jain, R. (2014). Leveling transparency via situated, intermediary learning objectives. In J. L. Polman, E. A. Kyza, D. K. O'Neill, I. Tabak, W. R. Penuel, A. S. Jurow, K. O'Connor, et al. (Eds.), Learning and becoming in practice. In *Proceedings of Learning and Becoming in Practice, the 11th International Conference of the Learning Sciences (ICLS 2014)* (Vol. 1, pp. 23–30). Boulder, CO: International Society of the Learning Sciences.

Abrahamson, D., & Shulman, A. (2017). Constructing movement in mathematics and dance: An interdisciplinary pedagogical dialogue on subjectivity and awareness. Paper presented at the 1st annual meeting of Movement: Brain, Body, Cognition, Oxford, UK, July 9–11. Accessed 02 Nov 2017 at https://edrl.berkeley.edu/sites/default/files/2017-07/Abrahamson.Shulman.MOVEMENT2017.pdf.

Bakker, A., & Derry, J. (2011). Lessons from inferentialism for statistics education. In K. Makar & D. Ben-Zvi (Eds.), The role of context in developing students' reasoning about informal statistical inference [Special issue]. *Mathematical Thinking and Learning, 13*(1–2), 5–26.
Bamberger, J., & Schön, D. A. (1983). Learning as reflective conversation with materials: Notes from work in progress. *Art Education, 36*(2), 68–73.
Bamberger, J., & Schön, D. A. (1991). Learning as reflective conversation with materials. In F. Steier (Ed.), *Research and reflexivity* (pp. 186–209). London: SAGE Publications.
Bereiter, C. (1985). Towards the solution of the learning paradox. *Review of Educational Research, 55*(2), 210–226.
Brousseau, G. (1997). *Theory of didactical situations in mathematic*. Boston: Kluwer Academic Publishers.
Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General, 12*(4), 355–376.
Chase, K., & Abrahamson, D. (2013). Rethinking transparency: Constructing meaning in a physical and digital design for algebra. In J. P. Hourcade, E. A. Miller, & A. Egeland (Eds.), *Proceedings of the 12th Annual Interaction Design and Children Conference (IDC 2013)* (Vol. Short Papers, pp. 475–478). New York: The New School & Sesame Workshop.
Chase, K., & Abrahamson, D. (2015). Reverse-scaffolding algebra: Empirical evaluation of design architecture. *ZDM Mathematics Education, 47*(7), 1195–1209.
Clement, J. (2000). Analysis of clinical interviews: Foundations and model viability. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 547–589). Mahwah, NJ: Lawrence Erlbaum Associates.
Cobb, P., Confrey, J., diSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational Researcher, 32*(1), 9–13.
Cole, M., & Wertsch, J. V. (1996). Beyond the individual-social antinomy in discussions of Piaget and Vygotsky. *Human Development, 39*(5), 250–256.
Dickinson, P., & Eade, F. (2004). Using the number line to investigate the solving of linear equations. *For the Learning of Mathematics, 24*(2), 41–47.
diSessa, A. A. (2007). An interactional analysis of clinical interviewing. *Cognition and Instruction, 25*(4), 523–565.
Edelson, D. C. (2002). Design research: What we learn when we engage in design. *The Journal of the Learning Sciences, 11*(1), 105–121.
Freudenthal, H. (1971). Geometry between the devil and the deep sea. *Educational Studies in Mathematics, 3*(3/4), 413–435.
Ginsburg, H. P. (1997). *Entering the child's mind*. New York: Cambridge University Press.
Goldstein, I., & Papert, S. (1977). Artificial intelligence, language, and the study of knowledge. *Cognitive Science, 1*(1), 84–123.
Gravemeijer, K. P. E. (1999). How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning, 1*(2), 155–177.
Greeno, J. G. (1998). The situativity of knowing, learning, and research. *American Psychologist, 53*(1), 5–26.
Hancock, C. (1995). The medium and the curriculum: Reflections on transparent tools and tacit mathematics. In C. Hoyles, R. Noss, & A. A. diSessa (Eds.), *Computers and exploratory learning* (pp. 221–240). New York: Springer.
Jong, T. D., & Joolingen, W. R. V. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*(2), 179–201.
Jordan, B., & Henderson, A. (1995). Interaction analysis: Foundations and practice. *Journal of the Learning Sciences, 4*(1), 39–103.
Kamii, C. K., & DeClark, G. (1985). *Young children reinvent arithmetic: Implications of Piaget's theory*. New York: Teachers College Press.
Kamii, C. K., & Dominick, A. (1998). The harmful effects of algorithms in grades 1-4. In L. J. Morrow & M. J. Kenney (Eds.), *The teaching and learning of algorithms in school mathematics, 1998 yearbook* (pp. 130–140). Reston, VA: NCTM.
Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science, 38*(6), 523–550.
Karmiloff-Smith, A., & Inhelder, B. (1975). If you want to get ahead, get a theory. *Cognition, 3*(3), 195–212.
Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.
Kirsh, D. (1996). Adapting the environment instead of oneself. *Adaptive Behavior, 4*(3–4), 415–452.

Kirsh, D. (2009) Projection, problem space and anchors. In: TN H, van Rijn, & L. Schomaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2310–2315). Mahwah, NJ: Lawrence Erlbaum.

Kirsh, D. (2010). Thinking with external representations. *AI & Society, 25,* 441–454.

Klahr, D., & Nigam, M. (2004). The equivalence of learning paths in early science instruction effects of direct instruction and discovery learning. *Psychological Science, 15*(10), 661–667.

Kuhn, D. (1995). Microgenetic study of change: What has it told us? *Psychological Science, 6*(3), 133–139.

Lave, J. (1992). Word problems: A microcosm of theories of learning. In P. Light & G. Butterworth (Eds.), *Context and cognition: Ways of learning and knowing* (pp. 74–92). Hillsdale, NJ: Lawrence Erlbaum Associates.

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science, 29*(4), 587–625.

Meira, L. (1998). Making sense of instructional devices: The emergence of transparency in mathematical activity. *Journal for Research in Mathematics Education, 29*(2), 129–142.

Meira, L. (2002). Mathematical representations as systems of notations-in-use. In K. Gravemeijer, R. Lehrer, B. V. Oers, & L. Verschaffel (Eds.), *Symbolizing, modeling and tool use in mathematics education* (pp. 87–104). Dordrecht, The Netherlands: Kluwer.

Noss, R., & Hoyles, C. (1996). *Windows on mathematical meanings: Learning cultures and computers*. Dordrecht: Kluwer.

Papert, S. (1980). *Mindstorms: Children, computers, and powerful ideas*. NY: Basic Books.

Parnafes, O., & diSessa, A. A. (2013). Microgenetic learning analysis: A methodology for studying knowledge in transition. *Human Development, 56*(1), 5–37.

Piaget, J. (1965). *The moral judgment of the child.* New York: The Free Press.

Piaget, J. (1968). *Genetic epistemology*. New York: Columbia University Press.

Pirie, S. E. B., & Kieren, T. E. (1994). Growth in mathematical understanding: How can we characterize it and how can we represent it? *Educational Studies in Mathematics, 26*(2–3), 165–190.

Roll, I., Aleven, V., Koedinger, K. R. (2011). Outcomes and mechanisms of transfer in invention activities. In: L Carlson (Eds.) *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Boston, MA: Publisher.

Ryle, G. (1945). Knowing how and knowing that: The presidential address. *Proceedings of the Aristotelian Society, 46,* 1–16.

Schön, D. A. (1992). Designing as reflective conversation with the materials of a design situation. *Research in Engineering Design, 3,* 131–147.

Schwartz, D. L., Chase, C. C., Oppezzo, M. A., & Chin, D. B. (2011). Practicing versus inventing with contrasting cases: the effects of telling first on learning and transfer. *Journal of Educational Psychology, 103*(4), 759–775.

Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist, 46*(6), 606–620.

Vergnaud, G. (1994). Multiplicative conceptual field: What and why? In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 41–59). Albany, NY: SUNY Press.

Vergnaud, G. (2009). The theory of conceptual fields. In T. Nunes (Ed.), Giving meaning to mathematical signs: psychological, pedagogical and cultural processes. *Human Development [Special Issue], 52,* 83–94.

von Glasersfeld, E. (1992). *Aspects of radical constructivism and its educational recommendations (Working Group #4).* Paper presented at the Seventh International Congress on Mathematics Education (ICME7), Quebec.

Walkington, C., Petrosino, A., & Sherman, M. (2013). Supporting algebraic reasoning through personalized story scenarios: How situational understanding mediates performance. *Mathematical Thinking and Learning, 15*(2), 89–120.

Zohar, R., Bagno, E., Eylon, B.-S., & Abrahamson, D. (2017). Motor skills, creativity, and cognition in learning physics concepts. In *Paper presented at the 1st annual meeting of Movement: Brain, Body, Cognition, July 9–11*. UK: Oxford