International Journal *of* Designs for Learning
2016 | Volume 7, Issue 2 | Pages 74-92

THE INVENTION COACH: INTEGRATING DATA AND THEORY IN THE DESIGN OF AN EXPLORATORY LEARNING ENVIRONMENT

Jenna Marks[1], Deena Bernett[2], & Catherine C. Chase[1]
[1]*Teachers College, Columbia University;* [2]*University of Chicago*

This design case describes the development of the Invention Coach, an intelligent exploratory learning environment (ELE) for the constructivist activity of Invention. The Invention Coach scaffolds students as they invent mathematical formulas to describe contrasting cases. In this paper, we detail our process and rationale for three key design decisions made with the goal of providing optimal guidance for Invention: (a) engaging in systematic analysis of teachers guiding students as a starting point for development; (b) developing a style of guidance that problematizes students' work through constraints, contrasts, and prompts for explanation; and (c) structuring the space through a nonlinear, modular set of activities focused on problem subgoals. We revisit the age-old assistance dilemma, discussing the unique difficulties of designing guidance and support in a computer environment for exploratory learning. Throughout, we discuss the tensions between providing adequate guidance and encouraging student exploration. We end by considering the trade-offs and unforeseen challenges that have come out of our design.

Jenna Marks is a Ph.D. student in Cognitive Studies in Education at Teachers College, Columbia University. She holds an M.A. in Instructional Technology and Media from Teachers College, Columbia University. Her research interests include exploratory learning environments, the motivational benefits of design thinking, and persistence in the face of failure.

Deena Bernett is a Ph.D. student in Developmental Psychology at the University of Chicago. She holds an M.A. in Cognitive Studies in Education from Teachers College, Columbia University. Her research involves the development of effective instructional interventions to support flexible understanding and transfer in mathematics.

Catherine C. Chase is an Assistant Professor of Cognitive Studies in Education at Teachers College, Columbia University. She holds a Ph.D. in Learning Science and Technology Design from Stanford University's School of Education. She studies how exploratory learning activities can be designed and supported to promote deep learning, transfer, and motivation in STEM domains.

INTRODUCTION

Computer-supported constructivist learning environments, such as those for exploratory learning and inquiry pedagogy, call for instructional designs that provide support and foster student-centered learning in ill-defined problem spaces. This design case presents the Invention Coach, an intelligent exploratory learning environment (ELE) to support middle school students through the process of Invention. Invention is a constructivist activity in which students invent a deep principle or structure to describe a set of contrasting cases—examples that differ on key features—calling students' attention to critical components of the underlying structure (Schwartz & Martin, 2004). Prior research has demonstrated that paper-based Invention followed by traditional expositions leads to robust learning and transfer (Roll, Aleven, & Koedinger, 2011; Schwartz, Chase, Oppezzo, & Chin, 2011; Schwartz & Martin, 2004). However, students reach frequent impasses during Invention tasks and require guidance and support to productively persist.

In developing the Invention Coach, we face the pedagogical tension of designing software-based scaffolds for exploratory learning. The challenge is to guide students towards productive interactions without being overly prescriptive, lest we compromise the process of discovery fundamental to this pedagogy. This tension is more broadly referred to as the *assistance dilemma* (Koedinger & Aleven, 2007); when, how much, and what kind of assistance should a learning environment provide to optimize student learning?

This design case details our process and the rationale for design decisions. Through a synthesis of exploratory research, prior work, and theory, we are working to develop an ELE that fosters productive exploration while providing adequate support. We focus on three key decisions made with this goal of providing optimal guidance for Invention: (a) engaging in systematic analysis of human teachers guiding students as a starting point for development; (b) developing a style of guidance that problematizes students' work through constraints, contrasts, and prompts for explanation; and (c) structuring the space through a nonlinear, modular set of activities focused on problem subgoals.

What Is Invention?

Invention is an instructional technique that is highly effective in cultivating robust understanding and transfer of deep structures found in science and math, such as variance, ratio, or vector components (Schwartz et al., 2011; Schwartz & Martin, 2004; Shemwell, Chase, & Schwartz, 2015). Invention begins with a period of exploration where students are asked to "invent" representations of deep structures (often mathematical formulas or physical principles) from a set of examples. During the process of Invention, students develop a differentiated understanding of the target domain, often noticing critical features and experimenting with ways to relate them into a coherent structure. *After* this initial period of exploration, students are told the canonical structures and formulas through traditional modes of exposition (e.g., lecture, reading). In this way, Invention differs from conventional "tell-and-practice" pedagogies, where students are typically told the formulas and structures *first* and are then asked to practice applying them in a series of problems. Several studies have demonstrated that compared to traditional tell-and-practice instruction, Invention is more effective at promoting deep learning and especially transfer of learning to novel contexts (Roll, Aleven, & Koedinger, 2009; Roll et al., 2011; Schwartz & Bransford, 1998; Schwartz et al., 2011; Schwartz & Martin, 2004).

Student inventions must accurately describe a set of "contrasting cases" or examples that differ on key features, which provide subtle guidance during the Invention process. Inviting students to generate inventions that work for all cases pushes students to identify critical features of the domain and combine them in a meaningful way. For example, in one study students were provided with four "pitching machine" cases, depicted as pitching target grids (Schwartz & Martin, 2004; see Figure 1). Tasked to invent a reliability index for "pitching," students compared these cases to notice important features, such as the spread of the dots and sample size, and they attempted to come up with a method for explaining the variance in the dots. Unbeknownst to the students, they were working towards inventing the formula for mean deviation ($\frac{\sum|x-\bar{X}|}{n}$), a simple way to compute variance.

Importantly, students need not come up with the correct answer during the Invention phase to learn from Invention activities (Ha & Sears, 2012; Kapur, 2010; Loibl & Rummel, 2014). Instead, researchers have suggested that Invention is effective because it supports productive exploration and creates "a time for telling" by preparing students to later appreciate the utility of the correct, formal equation (Bransford, Franks, Vye, & Sherwood, 1989; Schwartz & Bransford, 1998; Schwartz & Martin, 2004). For example, students tasked with inventing a "pitching index" for the cases in Figure 1 rarely generate the correct formula. However, through Invention they develop some prior knowledge of the critical components of calculating variance (e.g., sample size, spread of the data, the central tendency of the distribution), which prepares them to learn from the subsequent lecture explaining the canonical formula. Through Invention, students engage meaningfully with the cases, and, regardless of whether or not they invent a correct answer, they are prepared to learn the formal equation and understand its value.

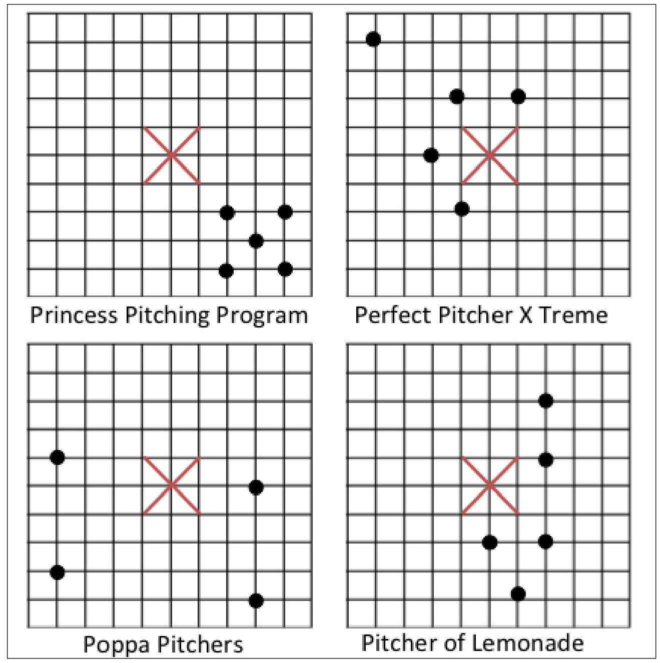

FIGURE 1. "Pitching" Invention Task (adapted from Schwartz & Martin, 2004). The goal is to invent a "reliability index" for each pitching machine. Each grid represents throws from a different machine, the "X" represents the target, and the dots represent where pitches land.

The Crowded Clowns Task

Our ELE is based on a previously designed Invention task, the Crowded Clowns task, in which students are asked to invent a numerical "index" to describe how crowded a set of clowns are in a set of buses (Schwartz et al., 2011; see Figure 2). Clown crowdedness serves as a proxy for density, where density is conceived as a number of objects crowded into a space (number of clowns / number of bus spaces). On a more abstract level, this Invention task prepares students to learn about ratio structures, elucidating the way the two critical features have opposite effects on an outcome. In this

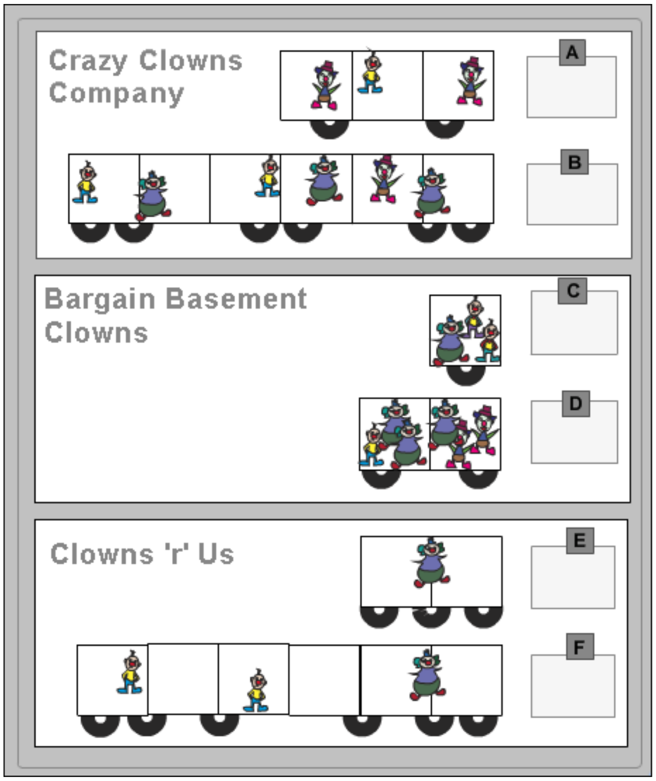

FIGURE 2. Crowded Clowns Invention Task (adapted from Schwartz et al., 2011). The goal is to invent an index of "clown crowdedness" for each bus. The task includes six buses distributed across three clown companies. Each pair of buses from a company is equally crowded.

example, the crowdedness of a given bus depends on both the number of clowns and the size of the bus, in an inverse relationship. Students are provided with the contrasting cases in Figure 2 and a set of four rules or problem constraints their inventions must adhere to:

1. You must come up with one number to stand for each bus' crowded clown index. The index will show exactly how crowded the clowns are.
2. A big index number means the clowns are more crowded, and a small index number means the clowns are less crowded.
3. You have to use the exact same method to find the index for each bus.
4. Buses from the same company are equally crowded, so they should have the same index.

Most students will initially attempt to describe crowdedness using a single feature—the number of clowns. However, the contrasting cases can help students notice the size of the bus, an important quantitative feature they often neglect to notice, by accentuating key differences that systematically vary across cases. For example, Cases A, C, and F vary on size but have the same number of clowns. Contrasting these buses should help students notice the importance of bus size, a less salient feature.

When attempting this task, students often generate many failed solutions to the Crowded Clowns problem as they gradually come to realize that a workable solution must involve both critical features in an inverse relationship. While many students do not produce the correct formula, the goal of Invention is productive exploration, which prepares students to learn from the later lecture on ratio structure.

Rationale for Invention Coach

The Crowded Clowns problem and other similar tasks have been used in classroom studies that suggest Invention is a highly effective pedagogy, particularly for helping students notice and apply deep structures in novel contexts. However, students require extensive help to work through Invention tasks productively. In particular, we have noticed anecdotally in studies of paper-based Invention that students have difficulty understanding the task and generating new ideas when old solutions fail. Classroom studies of Invention have required roughly a 1:5 teacher to student ratio in order to provide adequate help, but most middle school classrooms contain over 20 students with a single teacher. By building an ELE for Invention, we can make the pedagogy practical for classrooms by allowing adequate guidance to each student. Moreover, a computerized system to guide Invention can scale to enable this effective pedagogy to reach more students. This motivated the design of the current Invention Coach.

DESIGNING GUIDANCE FOR EXPLORATORY LEARNING ENVIRONMENTS

In designing our ELE for Invention, we aim to determine the optimal amount, type, and timing of support for students. This assistance dilemma, or the set of questions around correct parameters for effective feedback, has primarily gained traction in the world of Intelligent Tutoring Systems (ITSs). ITSs typically support students with step-by-step feedback and hints, guiding them through a well-defined solution path to learn specific skills and content (VanLehn, 2006). While numerous studies have demonstrated a need for significant pedagogical support for students in exploration activities (Geier et al., 2008; Mayer, 2004; van Joolingen, de Jong, Lazonder, Savelsbergh, & Manlove, 2005), how to scaffold exploration with adaptive support is a fairly new research question (Mavrikis, Gutierrez-Santos, Geraniou, & Noss, 2013). In contrast with ITSs, ELEs often have ill-defined solution paths, where multiple approaches to the problem space may be successful. Furthermore, exploratory learning is constructivist; it prescribes active, student-centered pedagogy. Therefore, the amount, type, and timing of support given for exploratory learning will differ in many ways from that in traditional ITSs. Instructional designers of ELEs have taken on the assistance dilemma in a variety of ways, summarized in the following section.

Scaffolds for Exploratory Learning

Cognitive scaffolds are tools that support students in solving problems that are otherwise too difficult. Prior work on scaffolding for ELEs has suggested a common tension: how can we provide adequate support for a student without reducing the inquiry process to "following cookbook instructions" (van Joolingen, de Jong, & Dimitrakopoulou, 2007, p. 112)? While students require help in order to learn effectively in ELEs, providing too much help can destroy the exploratory nature of the task. Therefore, we must strike a careful balance in designing support for ELEs.

On one side of the spectrum, traditional scaffolds directly support learning by structuring the problem space (de Jong, 2006; Wood, Bruner, & Ross, 1976). These structuring scaffolds decompose a complex task and restrict the problem space at opportune moments (Reiser, 2004; Wood et al., 1976). For instance, the one existing ELE for Invention, The Invention Lab, is designed as a linear set of three process steps that students are moved through sequentially as they invent structures for statistics (Roll, Aleven, & Koedinger, 2010). Many other ELEs similarly include explicit sequences of activities (Linn, 1995; Linn, Clark, & Slotta, 2003; Mavrikis et al., 2013). Another way to structure is to provide supports and spaces that help students construct arguments, develop explanations, and evaluate results. Structured hypothesis scratchpads, used in a variety of ELEs including SimQuest and Co-Lab, are templates that students fill in with relations and variables (de Jong, van Joolingen, & van der Meij, 2004; Mavrikis et al., 2013; Sandoval & Reiser, 1997; van Joolingen & de Jong, 1991; van Joolingen & de Jong, 2003).

In addition to traditional structuring scaffolds, an alternative promising strategy for scaffolding exploratory learning is by *problematizing* students' understanding. While scaffolds that structure the task *reduce* complexity, scaffolds that problematize subject matter *increase* complexity by focusing students on aspects of their work that are problematic. Problematizing occurs by a process of focusing a student's attention on an aspect of a situation and then creating a sense of cognitive dissonance by (a) illustrating how the student's understanding is flawed or (b) challenging his or her current understanding by requesting further clarification (Festinger, 1957; Reiser, 2004). Other researchers frame this process as introducing cognitive conflict or counter-examples (Mavrikis et al., 2013) or highlighting discrepancies between what a student has produced and correct productions (Wood et al., 1976). Many tools and scaffolds within ELEs that structure can also be leveraged towards problematizing. For instance, tools that elicit articulation and decision-making, particularly those requiring explicit explanations with constrained options, can make gaps in understanding transparent.

In summary, support for ELEs is achieved through two types of scaffolds: those that structure and those that problematize. As designers, we need to achieve a careful balance between both. Too much structuring can wreck the exploratory nature of the task. Too much problematizing can leave a student floundering. This general framework is embedded in our Invention Coach design, as we sought to develop structuring and problematizing scaffolds specific to the Invention task. In this paper, we focus on our development of problematizing scaffolds, as these are less common in computer-supported learning environments.

How to Support Invention

What are productive scaffolds for Invention, specifically? While several studies have explored the outcomes of Invention—success on Invention tasks and performance on tests of learning and transfer—what goes on *while* students are inventing, particularly at a fine-grained level of analysis, is largely a black box. We know of no study of moment-to-moment student actions and progress during Invention. Moreover, no studies have explored fine-grained student-teacher interactions as students are guided through the Invention process. Without careful studies of the process of Invention and how it might be guided on a moment-to-moment level, it is difficult to design an ELE for Invention. This led us to run our own exploratory study to observe naturalistic teacher guidance of the Invention process, which we describe later in the paper.

Despite the lack of fine-grained research, prior research and theory have identified some global-level guidelines for scaffolding Invention. For instance, prior studies have found that success on Invention and similar tasks is *not* related to learning and transfer (Kapur, 2010; Loibl & Rummel, 2014). Therefore, support for Invention tasks should not be solution-oriented; rather than pushing students towards the correct answer, as many ITSs do, Invention support is aimed to promote productive exploration. That is, support is intended to help students construct understanding by exploring the cases and producing inventions such that they are prepared to learn from a later lecture (Schwartz & Martin, 2004).

While it is unclear from current literature what exactly "productive exploration" entails, research has identified several core components of Invention. First, all Invention paradigms withhold explicitly telling students the correct method until the end in order to avoid cutting short the exploration process (Schwartz et al., 2011; Schwartz & Martin, 2004). Next, both research and general intuition would indicate that the "invention" piece is important; generation of new ideas is a key part of the process (Roll et al., 2009). Additionally, several studies converge upon the importance of noticing and differentiating the key features of the problem (Ha & Sears, 2012; Schwartz, Branford, & Sears, 2005). As explained earlier, the contrasting cases are developed with this in mind: certain contrasts make various features salient. Other research has emphasized that students must begin to combine features into a mathematical formula to successfully learn and transfer (Ha & Sears, 2012). This move towards mathematical representations prepares students to appreciate the correct mathematical formula. Last, researchers suggest that Invention is powerful because it encourages metacognitive monitoring, pushing students to uncover knowledge gaps that are later filled in by expository instruction (Loibl & Rummel, 2014). These core components of productive exploration—noticing key features, relating features mathematically, generating many possible solutions, and identifying gaps in understanding along the way—informed the design of the pedagogical style we employ in the Invention Coach.

Roll et al.'s Invention Lab (2010) supports the Invention process through a linear set of steps including (a) ranking the cases, (b) expressing a method in equation form, and (c) evaluating and revising the method. These steps help students notice features, generate inventions, and uncover gaps, respectively. While our Invention Coach is similar in some respects, it is unique in that we have modeled our ELE on guidance from a human teacher. As will become clear in the description of our design, this has resulted in a system without sequential steps that is more interactive and guided, where students do far more work focusing on the features and structures of the problem.

THE PRESENT DESIGN CASE

Design Process

We are following a multi-phase approach, blending formal empirical research with iterative cycles of design and informal user testing. The process began with a study of one-on-one teacher guidance of paper-based Invention, described in greater detail in the next section. We analyzed the data from this study and combined it with prior research and theory to develop a style of guidance and process model of Invention, which informed the design of an initial prototype Invention Coach. Through user testing with middle school students, we have refined our design and built the current Invention Coach prototype, discussed in this paper. This prototype operates in Wizard-of-Oz (WOz) mode. All feedback and guidance is built into the system, but an experimenter or "Wizard" on a paired computer must select *when* to give *which* kind of guidance, based on his or her assessment of the student's current knowledge state.

Using the current WOz prototype, we recently conducted a second formal empirical study to assess the difference between two versions of the Invention Coach—a *full* version with all guidance described in this paper and a *minimal* version with hint-style guidance that excludes the module-style guidance described in this paper. There is precedence for WOz studies in developing pedagogical systems for exploratory learning (Mavrikis & Gutierrez-Santos, 2010). We have not yet analyzed the results of this study but include in this paper experimenter observations that inform our discussion of the design decisions we have made. In the next year, we aim to analyze the data from the WOz study and combine it with theory surrounding Invention and our own pedagogical expertise to develop a fully functional, adaptive Invention Coach.

Design Team and Context

Our Learning Sciences lab at Teachers College, Columbia University is developing the Invention Coach in partnership with Carnegie Mellon University's Human-Computer Interaction Institute (HCII), as part of a three-year research study funded by a National Science Foundation (NSF) Cyberlearning Grant. In addition to developing an adaptive ELE for Invention, the goals of the study are to understand the process of Invention and identify productive types of support for Invention activities.

The Invention Coach was designed by a team of three Teachers College researchers led by a Learning Sciences professor and advised by an HCII professor from Carnegie Mellon. We explicitly populated our design team with researchers who had experience teaching math and science, so they could use their pedagogical expertise to inform design decisions. The first author has been a part of this design team since its initiation and has been integrally involved in all steps of the design, including user testing and empirical work. She also serves as the liaison between the programming team authoring the software at Carnegie Mellon and the learning sciences team at Teachers College.

THREE KEY DESIGN DECISIONS

The following sections explicate three key design decisions born out of the unique challenges of developing an ELE, with a focus on Invention. First, we informed our design with an empirical analysis of one-on-one human tutoring. This is an uncommon method of understanding a task, but we feel it provided a unique and more naturalistic construction of what happens during Invention. Second, through a synthesis of our own research with prior work and theory, we developed a style of guidance that operates by problematizing student work (Reiser, 2004). We problematize by reminding students of problem constraints, pushing students to contrast the cases, and eliciting students' explanations of their inventions. Third, we structured our Invention Coach with the phases of the Invention process uncovered in our study, resulting in a flexible, modular system that supports students in five task subgoals. These three decisions had critical implications for our design. By modeling the Invention Coach off of real human teachers, scaffolding through problematizing strategies, and developing a modular structure of five phases, we have strived to design an ELE with optimal guidance to facilitate productive exploration.

AN ELE MODELED ON GUIDANCE FROM REAL HUMAN TEACHERS

Our first design decision occurred in the development of the process itself. To the best of our knowledge, this is the only ELE for Invention modeled off of guidance from human teachers. We felt that only through observing and analyzing coached Invention in context could we fully unpack the process of Invention. This enabled us to develop a robust understanding of how students move through the process and of what types of guidance help students to learn and transfer from the activities. There is a precedent for modeling learning environments off of human tutors (Graesser, Person, Harter, & Tutoring Research Group, 2000; Lepper, Woolverton, Mumme, & Gurtner, 1993; Stevens & Collins, 1977). Moreover, researchers contrasting human tutors with ITSs suggest that tutoring systems are often too teacher-centered with a narrow range of strategies and can be overly direct, while human teachers offer more subtle feedback in more indirect ways (du Boulay & Luckin, 2001).

In many ways our design was informally influenced by anecdotes and general impressions of what went on in the student-teacher sessions, and these impressions are offered throughout this paper. Yet, notably, this empirical methodology goes beyond using human tutors as an inspiration for the design. As we explain throughout this design case, quantitative data from these studies directly informs design decisions. We took this approach because often the supports that lead to deep learning and transfer may not be obvious simply from watching tutoring sessions. The relative freedom students have in exploratory learning environments makes it difficult to model the task and predict the challenges students may have once interacting with the system (Mavrikis & Gutierrez-Santos, 2010). Moreover, previous studies of Invention find no relationship between success on the task and transfer, and they are inconclusive in regards to what specific progress relates to transfer (Kapur, 2010; Loibl & Rummel, 2014; Schwartz et al., 2011; Schwartz & Martin, 2004). A systematic approach to analyzing teacher-student interactions could reveal new insight into the optimal guidance of Invention.

A typical approach to the design of tutoring environments involves a cognitive task analysis, following one or two content experts or students as they complete a specific task (Gordon & Gill, 1997). However, in our study we gathered data from several teachers as they guided students through Invention tasks. We then systematically identified forms of guidance and related these forms of guidance to transfer outcomes.

Study Design and Aims

Three science teachers with more than five years of experience guided 18 seventh and eighth grade students through two Invention tasks, one of which was the "Crowded Clowns" task used in our Invention Coach. Teachers were instructed to not give away the target scientific structure or key features to students, but otherwise were told to guide naturally. The teachers worked one-on-one with students, with an experimenter filming the interaction. Students took a pretest and a posttest, with conceptual and transfer items assessing understanding of density, rate, and ratio structure. For a more thorough discussion of this study, please see Chase, Marks, Bernett, Bradley, and Aleven (2015).

The primary aim of this study was to determine *what* types of scaffolds and guidance to include in the ELE. Using discourse analysis methods, we identified the kinds of guidance teachers gave during the Invention activity. Then, we assessed which forms of guidance were productive by exploring which question types, explanations, and feedback were related to learning and transfer outcomes. Next, we identified particular subgoals or phases of the Invention process. Finally, we explored the paths that teachers and students took through the Invention space to determine the potential advantages of particular sequential patterns of activity that lead to learning and transfer. In the following sections, we discuss the results of our study and how they inspired the various components of support in the Invention Coach.

INSTRUCTIONAL STYLE

After running a naturalistic study with human tutors to gain insight and empirical understanding of guidance during Invention, the second critical design decision we made was designing the particular instructional strategies and supports for the Invention Coach. While our study of coached Invention informs our design in critical ways, it is impossible to rely entirely on empirical data from one small study of human teachers to design an entire instructional system. Therefore, we synthesized the results of our study with scaffolding strategies curated from prior work on Invention and, more broadly, research on and design of ELEs in general. In this section we discuss the pedagogical style uncovered in our study and the scaffolding strategies by which the components of our design are purported to work: contrast-based problematizing, constraint-based problematizing, and explanation-based problematizing.

Study Results: Pedagogical Style

Overall, we found that teachers provided a student-centered style of guidance, encouraging student talk and reflection. Students were very vocal; the ratio of teacher to student statements was roughly 1:1, which differs from standard tutoring, where the ratio is typically 3:1 (Chi, Siler, Jeong, Yamauchi, & Hausmann, 2001). When teachers did talk, they posed questions twice as often as they gave explanations (Chase, Marks, Bernett, Bradley, & Aleven, 2015). This rate of teacher questioning is far more frequent than in typical tutoring, where teachers give explanations more than three times as often as they ask questions (Chi et al., 2001).

Furthermore, the only significant predictor of transfer was the frequency of explanation. Teacher explanations were inversely related to transfer such that students who received more teacher explanations were less able to transfer their knowledge from the Invention activities to novel contexts. This finding holds when we control for students' prior knowledge and success on the Invention task, suggesting that teachers were not merely giving more explanations to students with low prior knowledge or poor performance. This result suggests that teacher explanations may have cut short students' exploration, ultimately hindering their ability to transfer (Chase, Marks, Bernett, Bradley, & Aleven, 2015; Schwartz et al., 2011).

Additionally, we found that teachers gave very little direct feedback (e.g., "That's correct," "You're right"). Only 7% of teacher statements contained explicit correctness feedback about something a student did or said (Chase, Marks, Bernett, Bradley, & Aleven, 2015). This also differs from standard tutoring; Chi et al. (2001) found roughly double that amount of feedback (15%). This provision of little direct feedback also departs significantly from the pedagogical philosophy of many intelligent tutoring systems, which immediately notify students when they make an error.

Collectively, these findings elucidate a pedagogical style, coined by VanLehn, Siler, Murray, Yamauchi, and Baggett (2003), of "ask more and tell less" (p. 246). Rather than simply telling students what to think and do, teachers pushed students to do the thinking themselves. Through frequent questioning, teachers guided students to generate their own understanding of the problem. Rather than giving direct feedback, teachers guided students to self-diagnose and correct their own misunderstandings. This student-centered style inspired the instructional strategies we developed and implemented in the Invention Coach.

Design Implications: Problematizing Strategies

While the style of guidance found in our study is different from that of standard tutoring, this is not all that surprising when considering the literature on Invention. As mentioned earlier, there is a fundamental "don't tell" rule-of-thumb in guiding Invention; students should not be told the answer until after they have had time to explore and grapple with the key structures of the problem (Schwartz et al., 2011). General work on coached problem solving suggests that allowing students to make errors is good for learning, and impasses are powerful moments that motivate students to construct understanding (Kapur, 2012; VanLehn et al., 2003). In fact, the "ask more and tell less" pedagogical philosophy is central to Invention, with the goal to keep students thinking and talking. It is clear from both the data and prior work that we should ask more, tell less, and avoid giving direct feedback. However, what questions *should* we ask? How can we help students self-diagnose their misunderstandings without telling them their inventions are right or wrong and providing lengthy explanations? How does one guide without telling?

Our review of the literature on ELEs suggests striking a balance between structuring the task and problematizing student understanding. Prior studies of Invention recommend supporting students in generating ideas, noticing and integrating features, and uncovering their own knowledge gaps (Loibl & Rummel, 2014; Roll et al., 2009; Roll et al., 2010; Schwartz, Branford, & Sears, 2005). Synthesizing this established research, we have developed three ways of problematizing students' understanding that help them to engage in the core learning processes of Invention. To help students notice features, we focus their attention on key contrasts. By comparing and contrasting cases, learners often come to notice key features of deep concepts that they would otherwise overlook (Bransford et al., 1989). To uncover gaps in students' understanding, we create cognitive dissonance, poking holes in students' reasoning (Festinger, 1957). We do this by pointing out problem constraints that students' indices are violating. Another way to uncover knowledge gaps is to elicit self-explanation, which has been shown to encourage metacognitive monitoring (Chi, Leeuw, Chiu, & LaVancher, 1994). Studies suggest that articulating ideas within a guided framework leads to deeper understanding and transfer and can stimulate idea generation (Aleven & Koedinger, 2002; Klahr & Dunbar, 1988; van Joolingen & de Jong, 1991). These "contrast-based," "constraint-based," and "explanation-based" problematizing strategies are powerful ways to scaffold exploratory learning that fit well within the philosophy of Invention.

Structuring and problematizing are complementary forces in a design, with the joint goal of cultivating productive exploration of a learning environment (Reiser, 2004). Therefore, many of our supports that problematize do so while simultaneously structuring the space. In the section describing the design of the Invention Coach, we will map these scaffolding strategies onto each specific design component in further detail.

MODULAR STRUCTURE SUPPORTING FIVE PHASES OF INVENTION

A third critical design decision we made in developing the Invention Coach was the modular structure of the system. Inspired by the phases of Invention uncovered in our study, we have designed various tools and lengthy interactions that target the subgoals students work towards during Invention. In this section, we describe the process model that emerged in our study and align each of these phases with the components of the Invention Coach.

Study Results: Phases of Invention

Through systematic coding of teacher and student talk, we uncovered a process model of coached Invention with five general phases. These phases are subgoals that teacher-student pairs tackled as they worked towards a solution: (a) Understand the Problem, (b) Notice Features and Structures, (c) Generate an Invention, (d) Evaluate an Invention, and (e) Calculate (see Table 1). In the "Understand the Problem" phase, pairs worked to clarify the task goal and reviewed the various constraints of the problem (e.g., buses from the same company are equally crowded). In "Notice Features and Structures," teachers guided students to notice critical features, develop a conceptual understanding of crowdedness, and grapple with the relationships between the key features and crowdedness. In the "Generate an Invention" phase, students generated numerical indices for one or more cases. In the "Evaluate an Invention" phase, teacher-student pairs worked to evaluate the correctness of a potential solution and reflect on why a solution may not work. Last, the "Calculate" phase involved counting features, simplifying fractions, and other numerical manipulation (e.g., multiplication). These phases overlap with the steps designed in Roll et al.'s Invention Lab as well as with prior research on Invention that suggests the importance of noticing features and generating mathematical relationships (Ha & Sears, 2012; Roll et al., 2010).

Students moved between these phases fluidly throughout the task, often passing through a particular phase multiple times. Overall, "Notice Features and Structures" was the most frequently visited phase, followed by "Generate an Invention," "Calculate," "Evaluate an Invention," and finally, "Understand the Problem," which occurred the least (see Table 1).

PHASE FROM EXPLORATORY STUDY (% OCCURRENCE)	DESCRIPTION
Understand the Problem (6%)	Explain or describe task goal and constraints
Notice Features and Structures (44%)	Notice key features of the underlying structure (e.g., number of objects, space)
Generate an Invention (20%)	Generate a solution (i.e., an index) for a case
Evaluate an Invention (12%)	Evaluate the correctness of a solution
Calculate (18%)	Simplify/manipulate fractions

TABLE 1. Process Model of Invention. Five general phases uncovered in our exploratory study. Percent (%) indicates frequency out of all phases.

Furthermore, we found that these phases did not occur in any standard order, and the order was largely dictated by teachers, not by students. While sessions with all three teachers most often began with "Understand the Problem" (63% of all sessions), each teacher had a unique pattern of coaching. Teacher A's sessions centered on noticing features and structures. She shuttled frequently between the notice and calculate phases but rarely visited other phases. Teacher B followed a more linear path through the process, starting by explaining the problem, then noticing features and trying a few calculations, before falling into repeated cycles of generating and evaluating inventions. Finally, Teacher C focused his sessions on producing inventions, with frequent visits to other phases in no particular pattern. He often started by asking students to generate an invention, followed by a series of generate-evaluate cycles, while occasionally toggling to noticing features, calculate, and understand the problem phases.

While there were striking differences in the paths teachers took through the Invention process, there were no differences in learning or transfer gains amongst teachers. On average, students made sizeable gains from pretest to posttest with all teachers. This finding suggests that there are multiple effective paths through the Invention space and there may be no single "best path."

Design Implications: Flexible Subgoal Structure

Informed by the process model uncovered in our study along with the precedent for structuring an exploratory problem space through process steps, we have structured the Invention Coach by creating components to support each process phase. However, breaking from the tradition of ELEs that lead students through a set of linear process steps, we were inspired by the fluid and flexible pattern of phases observed in our study to create a nonlinear process with the student-centered generation of Invention at its core. Specifically, the Invention Coach starts students on the main interface where they begin to invent an index of clown crowdedness. Intermittently, students move into various phases as they receive solicited and unsolicited guidance in the form of short hints or lengthy modules or use student-initiated tools and tabs to access information. After receiving or accessing the support, students continue to invent on their own. Therefore, *when* students go into a particular phase is determined by a combination of student initiative and the Wizard's assessment of each student's particular needs.

The phases decompose the complex task to scaffold students through the messy process of Invention. Each component essentially restricts the problem space by focusing students' attention on a particular task or particular cases. Importantly, while the modular phase format serves to structure the task and reduce complexity, many of the components themselves are geared towards problematizing student understanding, leading students to grapple with the important ideas in the domain.

THE INVENTION COACH

As illustrated in Figure 3, we have carefully designed components of the Invention Coach to support each process phase from our study. Components that support "Understand the Problem" include a three-minute instructional video and a rules tab where students can access the rules of the problem as needed. Components that support "Generate an Invention" include index generation spaces alongside various hints. To support "Notice Features and Structures," we developed two extensive interactive "modules," Ranking and Feature Contrast, described in detail later in this paper. To support "Evaluate an Invention" we included constraint-based hints and a notepad, and we developed a third module called "Tell Me How." Last, we support "Calculate" with a calculator tool readily available to the students.

Furthermore, specific components align with each of our problematizing strategies. In "Notice Features and Structures," the Ranking and Feature Contrast modules are contrast-based problematizers; these lengthy interaction components leverage the pedagogical utility of contrasting cases to help students notice overlooked features and integrate them into their understanding of crowdedness. The Feature Contrast module, along with the Tell Me How module from the "Evaluate an Invention" phase, further problematize by eliciting explicit articulation of student understanding. Last, the constraint-based hints problematize by poking holes in students' ideas, reminding them of rules that their Inventions violate, thus facilitating further evaluation. Notably, these problematizing components inevitably facilitate the "Generate an Invention" phase. Students often reach impasses when inventing; they get stuck on a particular idea (e.g., crowdedness = the number of clowns) and are unable to come up with a new invention. By illustrating how a student's current index is flawed, these components reveal gaps in student understanding or point to overlooked features, which help students in developing new Invention ideas.

Structuring scaffolds are not explicitly defined in Figure 3 because almost all components serve a structuring function; they orient students to relevant parts of the problem space and provide limited modes of expression, thereby reducing problem complexity. For further discussion of the alignment between process phases, instructional components, and pedagogical strategies, please see Chase, Marks, Bernett, and Aleven (2015).

Next, we describe the main interface and components of the Invention Coach, including the tools readily available to a student at the onset of the task. We then describe the

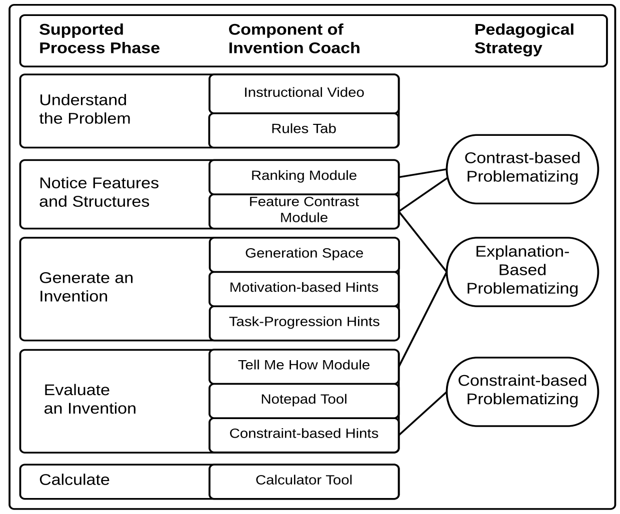

FIGURE 3. Alignment of process phases, Invention Coach components, and pedagogical strategies.

two levels of guidance built into the system: hints (short text guidance sent while the student attempts to invent an index) and modules (three extensive interaction sequences surrounding a particular subgoal). Throughout, we explain how components support their overarching process phases and align with specific scaffolding strategies.

Main Interface

The interface and its functions are explained to students through a short instructional video, where students are introduced to the task goal and constraints. As seen in Figure 4, the Invention Coach interface includes the contrasting cases (A), a predesigned element adapted from the paper-based version of the Crowded Clowns task (Schwartz et al., 2011). Students are able to select from five coach avatars prior to beginning the task, and the chosen Coach stands on the bottom right corner of the screen (B). Coach dialogue is delivered into a text box above the coach avatar (C).

Generation Space

Generation is the main goal of the Invention process. The student's task is to invent an "index" that explains how crowded the clowns are, and prior work on Invention suggests that coming up with new ideas is critical for learning and transfer (Kapur, 2012; Roll, et al., 2009). Moreover, the "Generate an Invention" phase accounts for 20% of coded episodes in our exploratory study. To support and structure this process phase, there are six input boxes to the right of the cases where students are able to generate and revise index numbers for each case (D). Students are readily able to generate inventions throughout the task, and the system is largely reactive to what students do in this space. All prompts, hints, modules, and tools appear on the right side of the screen and do not obscure the cases and index entry,

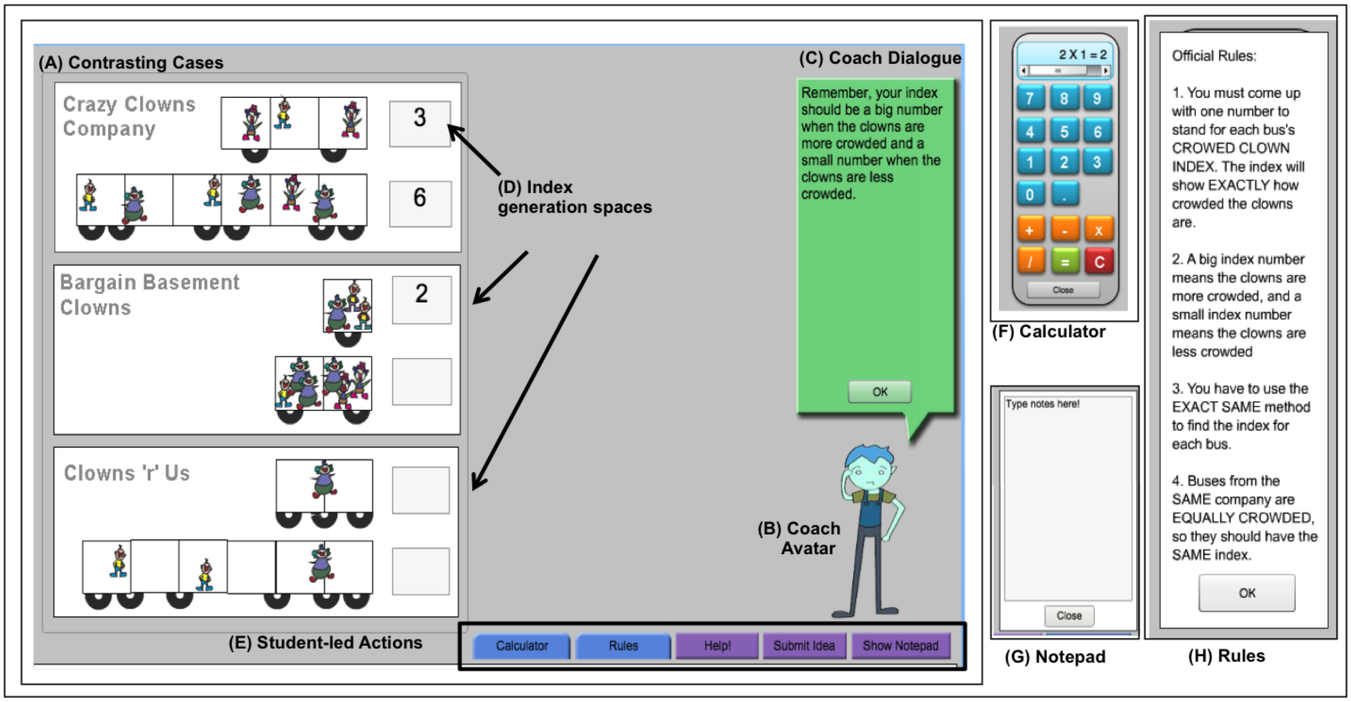

FIGURE 4. Invention Coach Main Interface and Student-Led Components.

allowing students to shuttle between provided scaffolding and the task at hand.

Student-Led Actions

To cultivate a student-led environment, we have built many tools that students can use autonomously within the system (E). Students can press "Submit" or "Help" to solicit guidance and have access to a calculator tool (F), a notepad (G), and a rules tab (H). In the instructions, it is emphasized that students are trying out ideas and can ask for help or hit submit at any time.

The right side of Figure 4 includes snapshots of these tools that would appear over the coach dialogue space in the interface. The calculator (F) offloads computational work so that students can focus on the goal of the task. It corresponds with the "Calculate" phase of our process model, which accounted for 18% of all episodes. Anecdotally, we found that many students in our study had great difficulty with simple division and fractions during the task. By automating this aspect of the task, the Invention Coach enables students to focus on the big ideas behind the math rather than getting tripped up on the nitty gritty details of calculation.

The notepad (G) was made available after user testing where many students requested paper and pencil. Here, students can jot down ideas that will not be evaluated by the system. This tool addresses the "Evaluate an Invention" phase, providing opportunity for self-monitoring by enabling students to express ideas in an unconstrained way.

The rules tab (H) structures the task by reiterating the task goal and providing explicit constraints for the student to follow in generating inventions. This addresses the "Understand the Problem" phase of our study, allowing students who are confused by the task to review the key goals and constraints. In paper-based Invention, students have the rules available at all times. In the Invention Coach, we go over the rules carefully during the instructional video and then keep the rules available via a clickable tab in order to simplify the screen space.

Hints

As shown in Figure 4, The Invention Coach provides short text guidance to the student in a dialogue box above the Coach avatar (C). Hints are given in both a solicited and unsolicited manner, and students are able to continue working on an Invention without addressing a hint. In our current WOz prototype, the Wizard can choose from a total of 25 hints. These hints are informed by both our study of coached Invention and prior work on scaffolding, and they were iteratively refined through user testing. While hints are currently sent by the Wizard as he or she feels necessary, we ultimately aim to develop "trigger conditions" for various hints in the final Invention Coach. The three categories of hints we developed are constraint-based hints, task-progression hints, and motivational hints.

Constraint-Based Hints

Constraint-based hints re-phrase specific rules and constraints of the Invention task that were provided in the instructions (e.g., "Don't forget. You have to use the exact same method to find the index for each bus"). These hints aid students in evaluating their inventions.

Constraint-based hints problematize by pointing out gaps in student understanding. For example, a student who is counting the number of clowns may recognize the flaw in his or her solution upon receiving a hint that highlights a constraint her solution violates (e.g., "Remember, your index should be a big number when the clowns are more crowded and a small number when the clowns are less crowded"). Students commonly have difficulty juggling the many inherent constraints to the task. In fact, Roll et al. (2010) relied on the constraints of the Invention problem to assess student answers in their ELE, providing feedback based on different violations of the rules. Of the three problematizing strategies built into the Invention Coach, the short constraint-based hints are the quickest method to illustrate gaps in student understanding.

Task-Progression and Motivational Hints

Task-progression hints structure the environment by reminding students of the affordances of the environment (e.g., "Try clicking on the rules tab to review the rules. That might help!") or focusing their attention on specific cases (e.g., "Okay. Look at the Clowns 'r' Us Company"). Motivational hints praise positive effort, giving vague positive progress information, and encouraging iteration (e.g., "I can see the gears turning in your brain," "You're thinking hard! Keep going"). These types of hints were common in our study of coached Invention. By giving students actionable steps to take and praising effort, these hints help students who are stuck or demotivated, therefore encouraging persistence and continued generation of ideas, supporting the "Generate an Invention" phase.

Modules

In addition to short hints, three modules in the Invention Coach engage students in lengthy interactions focused on problem subgoals. These modules were inspired by our Invention process model and borrow from prior research on Invention. The Ranking and Feature Contrast modules address the "Notice Features and Structures" phase derived from our study, and Tell Me How (TMH) addresses the "Evaluate an Invention" phase. While the modular format itself adds structure to the space, the main purpose of these modules is to problematize student understanding, leading students to grapple with the important ideas in the domain.

Like hints, modules are initiated by a Wizard, in any order, such that a student may do a module more than once or not at all. We aim to develop "trigger conditions" for each module to automate them in the final Invention Coach.

Ranking Module

The Ranking module was inspired by both a ranking activity in Roll's (2010) Invention Lab and observations of several ranking episodes in the teacher-student dialogue from our study. By inviting students to order the companies (pairs of cases) from most to least crowded, the Ranking Module compares students' index responses with their intuitive understanding of crowdedness (see Figure 5). After correctly ranking the cases, the module asks students to reflect on what led them to decide each bus' crowdedness. The Ranking module facilitates the "Notice Features and Structures" phase of Invention. It draws on the intuitive knowledge of crowdedness that students bring into the task, prompting students to recognize what aspects of the cases led them to intuitively rank in that particular order. By noticing and understanding these features, students are prompted to revise their inventions and develop increasingly sophisticated indexes that incorporate the critical features.

Ranking can serve as a structuring scaffold if used to orient a student when he or she is confused. However, Ranking also works as a contrast-based problematizer, helping students notice and understand critical features. When students compare their ranking to the indices they have chosen for each company, they can uncover gaps in understanding or notice new features that they can integrate into their idea of crowdedness. For example, many students will assign an Index of 1 to both cases in Crazy Clowns and cases in Clowns 'r' Us (see Figure 5). However, ranking pushes students to compare the cases and determine that Crazy Clowns are more crowded. This creates cognitive dissonance; both companies cannot

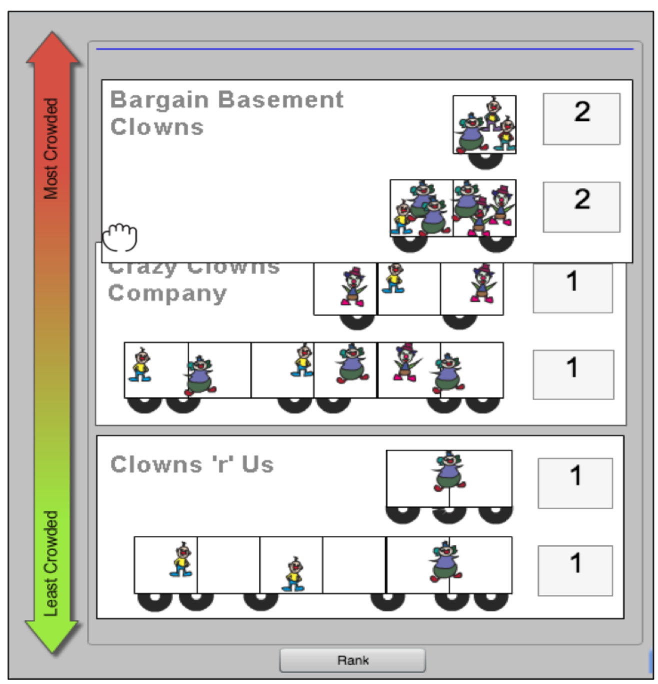

FIGURE 5. Ranking Module. Students drag-and-drop pairs of cases to order them from most to least crowded.

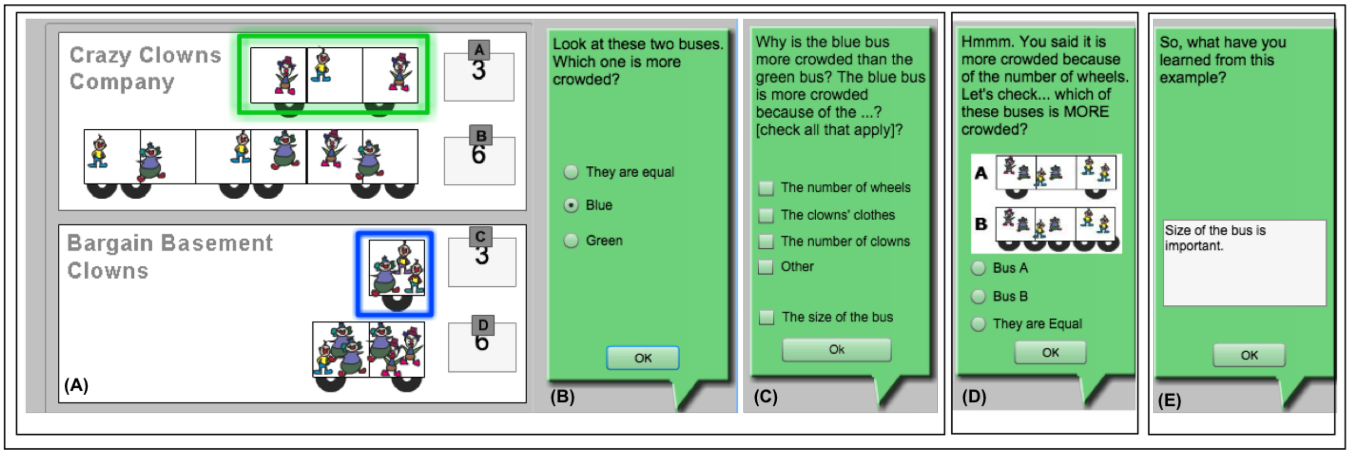

FIGURE 6. Feature Contrast Module. This module first highlights a key contrast of two buses (A), asks the student which bus is more crowded (B), invites the student to identify which feature causes one bus to be more crowded than another (C), problematizes misconceptions through additional contrasts (D), and elicits open-ended reflection on what the student has learned (E).

have the same index if one is more crowded than the other. This may lead students to think more deeply about what constitutes crowdedness.

Feature Contrast

To further support the "Notice Features and Structures" subgoal phase, we developed the Feature Contrast module. Feature Contrast also makes use of the contrasting cases to help students notice critical features and understand how they relate to crowdedness. However, the Feature Contrast module is much more focused than the Ranking module. Feature Contrast asks students to contrast two specific cases and pushes them to identify the features that contribute to crowdedness in them. There are six contrasts built into the Invention Coach, each consisting of two cases that are compared. Four hold the number of clowns constant, while bus size differs, enabling students to recognize the effect of the size of the bus (e.g., Figure 2, Cases B & D; A & C; A & F; C & F). Two hold the size of the bus constant, while the amount of clowns differs, enabling students to recognize the effect of the number of clowns (e.g., Figure 2, Cases B & F; D & E).

In this module, the Wizard selects one of the contrasts to "send" to the student. The two cases of interest are highlighted. For example, in Figure 6, the Wizard has highlighted Cases A and C for the student, to help him or her notice size of the bus and the way in which it affects crowdedness (A). The student is asked which of the two buses is more crowded (B). Students tend to have an intuitive understanding of crowdedness and usually correctly select the more crowded case. However, the student is then posed a second, more challenging question of *why*. Why is that case more crowded? Students must select from a "check all that apply" menu seeded with features (C). Comparing the two cases, students often notice the important feature that varies. Forcing students to make a decision from limited options encourages them to grapple with the critical concepts. In this example, the correct answer is the size of the bus (the feature that varies). However, many students err and choose the number of clowns, the more salient feature, or the number of wheels, an irrelevant feature. If they choose incorrectly, there are a variety of follow-up questions designed to contradict students' misconceptions and help them actively process errors. For example, if the student selects "the number of clowns", the follow-up question is "which of these buses has more clowns?" The answer (they are equal) problematizes the student's prior work. Or, if the student selects an irrelevant feature, he or she is presented with two new cases to compare (D). These cases are equally crowded but vary on the irrelevant feature only. The student is again asked which is more crowded, problematizing the student's understanding. Contrasts are designed to help students realize that certain features such as the number of wheels do not affect crowdedness while other, less salient features like bus size do. The module loops back to the original question until the student gets the answer correct. Once a student correctly chooses the "why," the student is given a free-text box to explain what he or she has learned from the contrast (E). Therefore, in addition to using contrasts, Feature Contrast problematizes through explanation. Feature Contrast compels students to explain first through a constrained response that highlights critical features and then through the free response where students can integrate the information gleaned from the module into their understanding of the problem. For instance, a student may write "size of the bus is important" in the text box (E).

Tell Me How

The second module that employs the explanation-based problematizing strategy is Tell Me How (TMH). This module induces articulation of and reflection on Invention ideas, supporting the "Evaluate an Invention" process phase. One of the key ways in which Invention is purported to work is by helping students identify gaps or flaws in their understanding (Loibl & Rummel, 2014). In addition to supporting students in the important process of articulation, TMH provides the system with much-needed diagnostic information about a student's current knowledge state.

Anecdotally, we noticed in our study of coached Invention that each time a student wrote down an index, the first question out of the teacher's mouth was often, "How did you get that index?" Therefore, similar to the hypothesis scratchpads described in our review of ELEs, this module asks students to "tell me how you got" a specific index. Students are restricted in how they can explain their methods, but the Invention Coach allows them to express understanding in multiple forms, facilitating connections between intuitive understanding and mathematical representations.

Importantly, the goal of TMH is not to lead students towards a correct answer or to give them explicit right/wrong feedback about their indexes. Recall that teachers in our study gave very little direct feedback, perhaps encouraging students to recognize their own errors. In keeping with this approach, the goals of TMH are to (a) encourage students to clearly articulate their solution process, which often leads students to self-diagnose errors and (b) help the system understand how students generated their solutions.

TMH includes a sentence maker, or cascading set of drop-down menus, that supports students in identifying their general problem approach. It then encourages students to link their index numbers to visual referents in the cases. It also facilitates expression of mathematical equations using the calculator tool. The terminology and options in the drop-down menus were designed by reviewing coaching session transcripts to best capture the ideas students had and the language they used to express these ideas. For example, we read through the transcripts to determine how most students described units of space (e.g., boxes, bus spaces, dividers, bus cars) to determine the language for that menu option.

The TMH module problematizes by enabling students to self-diagnose gaps in understanding when they articulate their methods. By tasking students with developing clear methods for Invention and connecting their inventions to specific referents in the cases, the TMH module drives students to grapple with how features relate to an index in well-defined ways. One tension in developing scaffolding tools is how to balance supporting intuitive strategies with requiring students to work within a disciplinary framework (Reiser, 2004). TMH moves students to more mathematical explanations, the ideal mode of expression within Invention.

In Figure 7, the student is asked to "Tell me how you got 1" as an answer for a specific case (A). The student can choose between the options "I counted," "I did math," or "I estimated." Choosing "I estimated" results in a prompt reminding the student that the goal of this task is to be exact and, therefore, not to estimate. If the student selects "I counted"

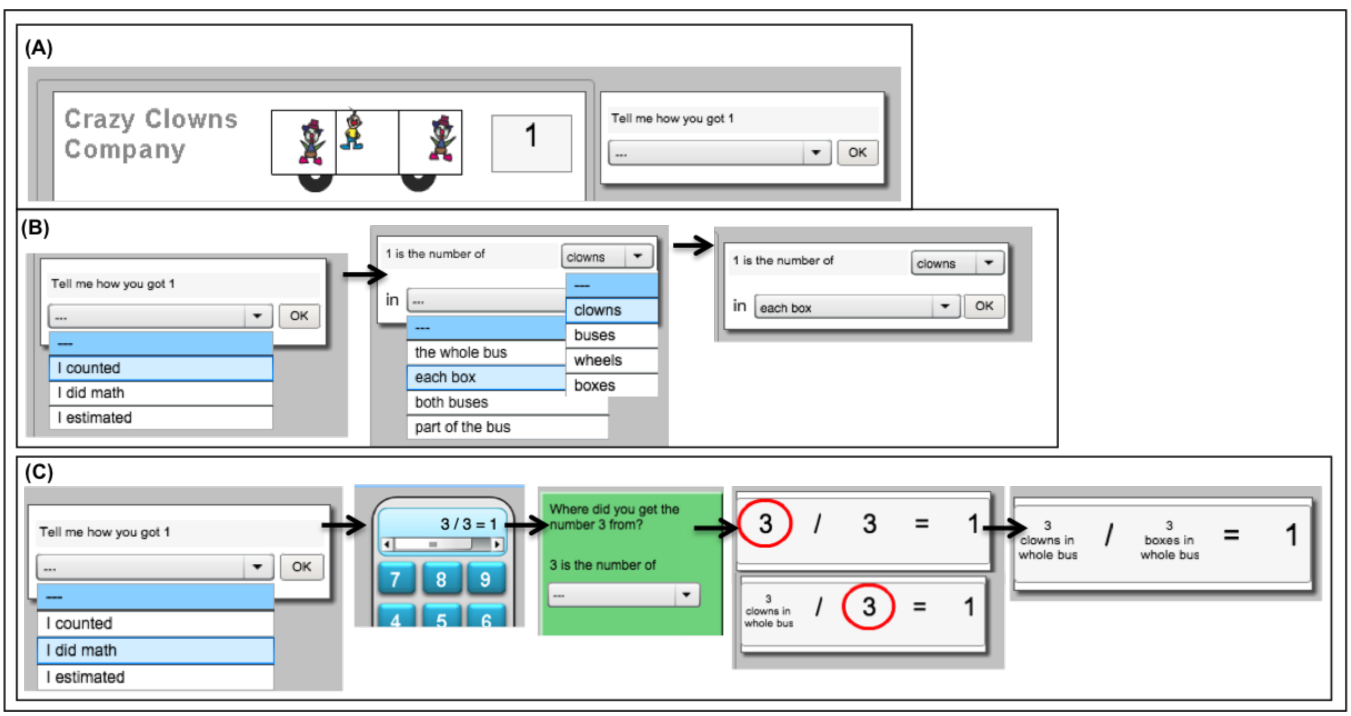

FIGURE 7. Tell Me How Module. This module invites the student to explain how she generated an answer for a specific case (A). If student selects "I counted," the student is asked to use a sentence maker to identify what was counted (B). If student selects "I did math," the student is provided with a calculator to input his or her equation, and then uses the sentence maker to explain the numbers in the equation (C).

(B), the student is brought to a sentence maker from which pre-seeded features can be selected to replace two variables that define what he or she counted; 1 is the number of (clowns/buses/wheels/boxes) in (the whole bus/each box/both buses/part of the bus). In this case, the student selects "clowns" and "each box". Having given a complete "counting" solution, the TMH module ends and the student continues the task.

If the student selects "I did math" (C), the student is invited to use the calculator to show the math that got him or her to that index. As can be seen in Figure 7, the student inputs "3 / 3 =" into the calculator. At this point, the equation is pasted in the TMH box, and the student is asked to label each number on the left side of the equation with the same sentence maker used for the "I counted" portion of the module.

Because TMH does not explicitly judge the correctness of a solution, the answer is satisfactory if the equation equals the index, even if it does not reflect the correct density equation. For example, if the student wrote "1 + 0 =" in the calculator for this case, this would also be a satisfactory response. If the student's equation does not equal the index, the student is prompted to change his or her index or change his or her math, leading the student to reflect on his or her incorrect reasoning. Being pushed to articulate their solution process often helps students uncover flaws in their indices, without requiring explicit right/wrong feedback from the system. For instance, a common wrong solution is for students to simply rank the cases (1, 2, 3) from least to most crowded. This would lead students to give both buses in the Crazy Clowns Company an index of 2, since this company is the second-most crowded. When asked to explain how they got 2, students might select "I estimated," to which the Coach suggests that the student rework their answer to be exact. If a student selects "I counted," the prompt to say exactly what two things they counted (clowns/buses/wheels/boxes) leads students to realize that they did not actually count out a precise solution. If students select "I did math," they are often flummoxed when the calculator appears and they realize they did not do an exact math calculation to arrive at their answer. In these situations, students typically decide to rework their solutions. Being forced to clearly delineate how they generated their answers can lead students to realize that a vague ranking solution is incorrect and that their solutions must be precise.

In addition to encouraging self-diagnosis of errors, the TMH module can transition students from an intuitive understanding to a mathematical one. There are two ways that students can express a correct understanding of crowdedness in the TMH module: a more intuitive one (I counted: clowns in each box) and a mathematical one (I did math: number of clowns in the whole bus / number of boxes in the whole bus). Allowing students to represent knowledge in either fashion, TMH bridges students' understanding to support sense making and articulation. In our exploratory study, we found that some students could intuitively see a certain number of clowns per box (producing a correct index for some cases), but they could not formulate this number mathematically, by constructing a ratio of clowns divided by boxes.

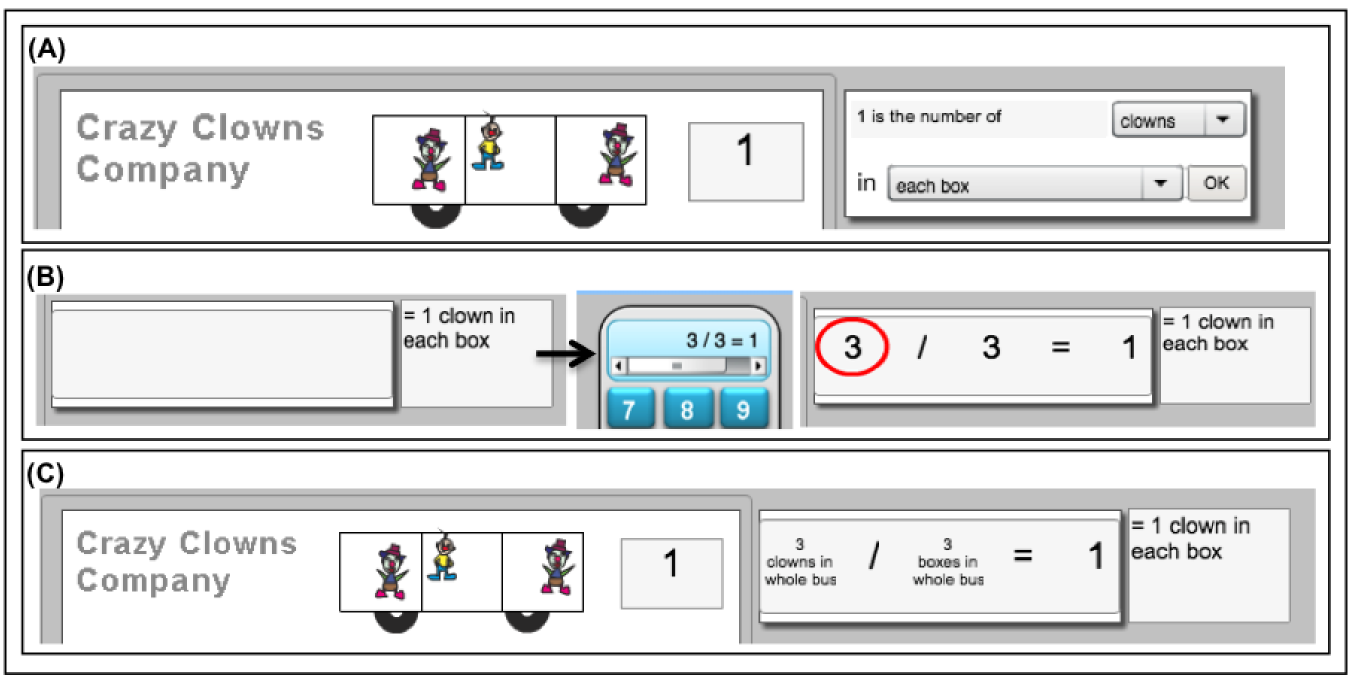

FIGURE 8. Transition to Math. This function takes a student's intuitive qualitative explanation of 1 clown per box (A), and tasks the student to come up with the same answer, using quantitative math (B). This ideally results in the student connecting an intuitive understanding of ratio to a mathematical one (C).

To help students transition from an intuitive understanding to a mathematical one, TMH includes a *Transition to Math* function (see Figure 8). This urges students to think about how they can transform their intuitive explanations (i.e., number of clowns in each box) to more mathematical ones (i.e., clowns in the whole bus / boxes in the whole bus). If a student has demonstrated a correct intuitive understanding for a specific index (A), this function moves the student's explanation to the right of the TMH box, clears the TMH box, and asks the student if he or she can come up with this same answer, using math (B). Ideally, the student will end up with the correct ratio explanation (C).

Sample Narrative of an Invention Process

To paint a clearer picture of how Invention works within our ELE, we describe the Invention process of "Victor" (name changed), an actual student who participated in the WOz study we recently conducted. In this session, a Wizard selected and executed various hints and modules, as she deemed necessary. Figure 9 is a sped up and edited down video of the student's screen during this fifteen-minute session. The student-facing interface does not include the A-E labels for these buses; we refer to the labeling visible on the wizard-facing screen shown in Figure 2.

Victor begins by inventing a common first index: counting the number of clowns in each bus, giving the first bus an index of 3, the second bus an index of 6, and so on. He clicks submit, and the Wizard gives Victor a Feature Contrast module of Cases B and D. This is designed to help Victor notice the importance of the size of the bus, the less salient feature that he has not used in determining his index. He is asked which of the two cases is more crowded and accurately chooses Case D. However, when asked why, he selects "the number of clowns." The system follows up by asking Victor which of the two buses has more clowns. Victor recognizes that they have an equal number of clowns. If these buses have the same number of clowns, the number of clowns cannot cause one bus to be more crowded. The Invention Coach repeats the question: "Why is Case D more crowded than Case B?" Victor now selects "the size of the bus," and, when prompted for explanation, states, "It does not matter how many clowns there is what matters is the length of the bus." To combat this overgeneralization, the Wizard gives Victor a second Feature Contrast module, this one designed to help him notice the importance of the number of clowns. At the end of this module, Victor writes "If the bus is the same size as the bus that we are comparing it to the amount of clowns matters if the bus is crowded or not." This response

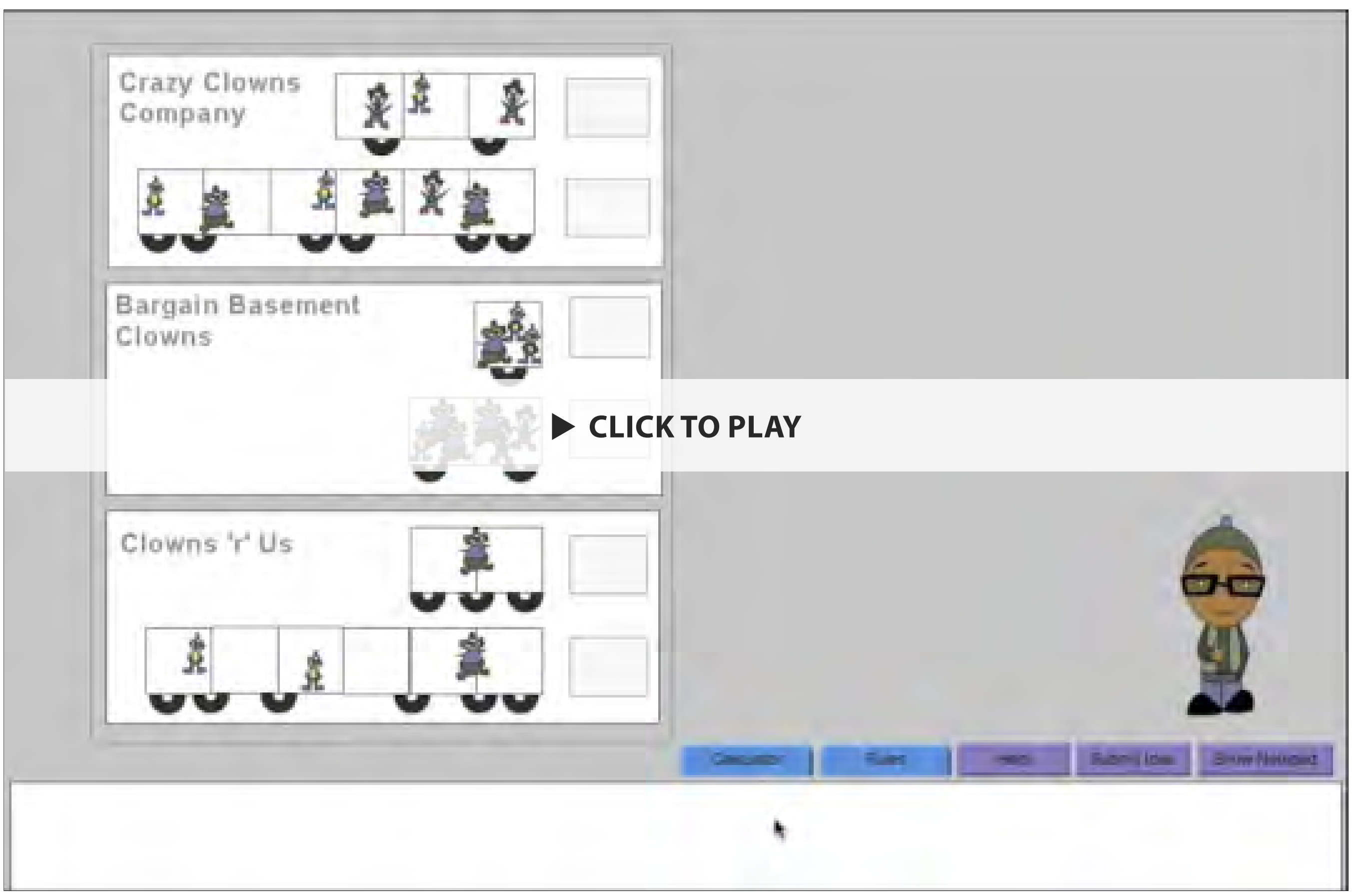

FIGURE 9. Video Narrative of Victor's Invention Process.

suggests preliminary reflection on how the two features interact to create crowdedness.

The Coach gives Victor motivational feedback: "It may not feel like it, but you're making good progress," and suggests he use what he just learned. Victor proceeds to change his indices and clicks submit. His pattern of responses is not clear; he gives Cases A, B, C, and D an index of 3 and Cases E and F an index of 1. To learn more about Victor's thinking and to encourage Victor to articulate his ideas, the Coach launches into a Tell Me How module for Case C, where he uses the drop down menus to express that "3 is the number of clowns in part of the bus." This is not correct (3 is actually the number of clowns in the whole bus), and Victor chooses to delete both indices C and D. He clicks "help."

To encourage Victor to use his intuitive understanding of crowdedness in developing his index, the Invention Coach launches a Ranking module. Victor correctly ranks the cases and then changes all of his index ideas. This time, his index is a count of the number of bus spaces in each bus (Case A is 3, Case B is 6, Case C is 1, Case D is 2, and so on). He clicks submit. The system gives him constraint-based feedback, reminding him that buses from the same company are equally crowded. His current response does not follow that rule. He clicks on the rules tab to review the rules, deletes his indices, and solicits more help.

The Coach sends a prompt to encourage idea generation: "It doesn't have to be perfect. Just try something!" Victor responds by trying another idea—this time giving both Crazy Clowns' buses an index of 2, both Bargain Basement Clowns' buses an index of 3, and both Clowns 'r' Us buses an index of 1—a solution that ranks the cases from least to most crowded. He clicks submit. The Invention Coach runs through one more Tell Me How module with Victor before the session ends due to time constraints.

While Victor did not develop a correct Invention, we must remember that success is not a requirement for transfer from the Invention task. From this session, it does indeed appear that Victor was productively exploring. He generated a number of different solution ideas, and he noticed and created Inventions that utilized both critical features (bus size and amount of clowns). This exploration was designed to prepare him to learn from the subsequent "tell": a video that explains the correct solution and teaches about ratio structures.

Pitfalls and Trade-offs in Our Current Design

We designed each component of the Invention Coach to scaffold students in productive exploration. We made numerous design decisions; many were empirically driven, others were inspired by prior work or theory, and for some decisions such as wording we relied on our own teaching expertise along with findings from user testing. While we have used good design practices such as data-driven design and iterative user testing, there are still many unresolved issues with our prototype. Some are ultimately trade-offs we have made and others we hope to address in our next iteration.

Too Much Structure?

In developing subgoal modules, we structure what is traditionally an open-ended problem space. For example, the trade-off in developing a highly structured articulation space such as TMH is that while it compels students to use the vernacular of the discipline and represent their knowledge in ways that the system can diagnose, it limits student expression and can be leading. By including the relevant features in the drop-down menus, we could be inadvertently tipping students off to their existence. Furthermore, students can use the drop-down menus to reverse engineer a method for an arbitrary number they put in a box.

In some ways, it seems as though we are leading students by the nose. In the video narrative, the Wizard helps Victor to notice bus size, and he responds by generating a solution using bus size alone. The Wizard asks him to rank, and he then develops a ranking solution. If the students are overly responsive to system feedback, is this truly exploration?

Additionally, many of these modules are laborious, taking a considerable amount of time, which takes valuable time away from the generation aspect of the task. In Ranking modules, while students generally have little trouble correctly ranking the cases, there are many text pop-ups before and after the ranking portion, explaining what to do in this module and asking students to reflect on their intuitive knowledge. From our observations in the WOz study, this module takes a disproportionately long amount of time to complete. Students similarly find themselves in lengthy Feature Contrasts if they continually answer incorrectly, leading to frustration and nonproductive exploration. Multiple students follow Victor's pattern of focusing only on one feature after a single Feature Contrast (the one it is designed to help them notice), requiring a second, lengthy Feature Contrast to demonstrate that both features are important.

How can we improve upon our design to avoid over-structuring? For our next iteration, we aim to simplify and shorten several modules. We also hope to craft an integrated Feature Contrast module that enables students to reconcile the importance of both features within a single execution of the module. Additionally, analysis of the log data from our WOz study could reveal which modules or components are most effective for learning and transfer, providing recommendations for structuring aspects that we could remove. By pruning support such that we only include scaffolds that have demonstrated effects on transfer, we can cultivate a more exploratory environment.

Ambiguous Feedback

An additional unforeseen problem with the current Invention Coach is the way in which some feedback has played out in the system. In an effort to limit direct feedback and foster problematizing through questioning, we have implemented a lot of ambiguous feedback, such as the text students see upon finishing a Tell Me How module: "Thanks for letting me know that. Let's keep going." However, observations from the WOz study indicate that students tended to misinterpret this feedback in unexpected ways, often interpreting vague responses from the Invention Coach as absolute correctness feedback about their inventions. For example, if a student accurately explained her index in a TMH, the student would often conclude that the index number was itself correct. Conversely, if a student had a correct index but could not explain it, feedback in TMH that stated the student's explanation did not work often resulted in the student believing the index itself was wrong. This feedback often stuck with the student for the remainder of his or her session. For example, a student input 3 for Case A and expressed it as "The number of clowns in the whole bus," which is factually true. TMH thanked her for explaining her idea and kept the idea on the screen. She audibly cheered for herself and never touched that 3 again, despite beginning to calculate other indices as "the number of clowns in each box." Similar incidents occurred in response to motivational or constraint-based hints, which often led students to conclude that all their work was either correct or incorrect, while the Invention Coach simply meant to praise student effort or problematize a single index.

The challenge of providing problematizing feedback in the Invention Coach highlights a broader challenge in the design of ELEs. While many intelligent tutoring systems immediately notify students when they make an error and explicitly highlight the error, ELEs often have the goal of getting students to diagnose errors on their own. In this case, indirect problematizing feedback was our solution but often led to incorrect interpretations by students.

These issues remind us that the Invention Coach does not exist in a vacuum. Students in today's educational climate are accustomed to being told when they are right and when they are wrong. It is likely that we did not see this problem in our exploratory study of coached Invention because human teachers have a greater ability to offer subtle feedback in more indirect ways, via tone, gesture, or body language (du Boulay & Luckin, 2001). This is one potential pitfall in using human interactions when designing ELEs. While the ELE can mimic teacher guidance, we do not have the same bandwidth for affect or overall communication that an in-person interaction affords.

Ultimately, we must consider classroom norms and the implicit assumptions students have around feedback when designing our Invention Coach. Therefore, we aim to make the intentions behind feedback and hints more transparent in our next iteration, potentially by explicitly telling students that we will not be telling them if they are right or wrong but simply trying to help them think about the problem in different ways. Particularly for TMH, we must develop language that makes it clear to a student that each time we give feedback, we are *not* necessarily telling the student whether he or she is correct or incorrect. Another promising addition for the next iteration is the inclusion of auditory feedback, which can recapture some of the affect that makes human tutors so effective with subtlety.

CONCLUSION

The Invention Coach is an exploratory learning environment that uses scaffolds to guide students through the challenging task of Invention, structuring the problem space and problematizing the inventions that students generate to uncover gaps in understanding. Overall, these scaffolds aim to facilitate productive exploration of the Invention space, which will lead to robust understanding and transfer of the target learning.

We developed the Invention Coach with an iterative cycle of design and a synthesis of exploratory empirical research, theory, and prior work on Invention. To our knowledge, this is the first ELE for Invention derived from systematic analysis of teachers working one-on-one with students. From our exploratory study of coached Invention, we uncovered an "ask more and tell less" instructional style and developed a process model of Invention with five phases. We discussed the ways in which two strategies of scaffolding—traditional structuring and less-traditional problematizing—can help students productively explore the environment. We adapted the strategy of problematizing to our ELE, explaining how our system problematizes student understanding in three overlapping ways by leveraging contrasting cases, highlighting constraint violations, and eliciting self-explanation. Finally, we mapped the phases and pedagogical strategies onto the actual components of our Invention Coach.

The Invention Coach is a work in progress, and we are excited to see how our design evolves. In our recent WOz study, 46 students tried two versions of the Invention Coach. We are eager to use this data to inform both refinements in Invention components and to determine "trigger moves" for guidance in order to develop a fully adaptive ELE.

Resolving the assistance dilemma within exploratory learning environments is a challenging undertaking, and here we present one method for doing so. We are still in the process of developing a fully functional prototype, and we continue to grapple with trade-offs. Do we structure the space in a way that helps us assess student understanding or provide more freedom to explore? Do we encourage self-diagnosis of misconceptions or give clearer feedback?

However, we see promise in using problematizing strategies, such as contrasting, constraining, and explaining, for exploratory learning. Ultimately, the Invention Coach design is a potential solution to the tension of under-assisting versus stifling Invention; we aim to create a system that feels less like "cookbook instructions" and more like improvising a meal from a set of ingredients, fostering structured yet constructivist exploration for students. We straddle the sandbox-to-scaffold spectrum, with one foot firmly rooted in supports that structure and the other in strategies such as problematizing that increase complexity and push students to explore, preparing them to learn from future instruction.

ACKNOWLEDGMENTS

The authors wish to thank Vincent Aleven and our team at Carnegie Mellon for their instrumental role in developing the Invention Coach. Our work is supported by the National Science Foundation under Grant No. 1361062. We thank the editors and reviewers for organizing this special issue.

REFERENCES

Aleven, V. A., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science, 26(2), 147-179.*

Bransford, J. D., Franks, J. J., Vye, N. J., & Sherwood, R. D. (1989). *New approaches to instruction: Because wisdom can't be told*. In S. Vosniadou & A. Ortony (Eds.), Similarity and analogical reasoning (pp. 470-497). New York, NY: Cambridge University Press.

Chase, C. C., Marks, J., Bernett, D., & Aleven, V. (2015). The design of an exploratory learning environment to support Invention. In J. Boticario & K. Muldner (Eds.), *Proceedings of the Workshops at the 17th International Conference on Artificial Intelligence in Education, AIED 2015,* Vol. 2 (pp. 1-8). Retrieved from http://ceur-ws.org/Vol-1432/iseole_pap1.pdf

Chase, C. C., Marks, J., Bernett, D., Bradley, M., & Aleven, V. (2015). Towards the development of an exploratory learning environment: A study of naturalistic teacher guidance of Invention. *Proceedings of the 17th International Conference on Artificial Intelligence in Education, AIED 2015,* (pp. 558-561). Amsterdam, NL: IOS Press.

Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science*, *18*(3), 439-477.

Chi, M. T., Siler, S. A., Jeong, H., Yamauchi, T., & Hausmann, R. G. (2001). Learning from human tutoring. *Cognitive Science*, *25*(4), 471-533.

de Jong, T. (2006). Scaffolds for scientific discovery learning. In J. Elen, & D. Clark (Eds.) *Handling complexity in learning environments: Research and theory* (pp. 107–128). London, UK: Elsevier Science.

de Jong, T., van Joolingen, W., & van der Meij, J. (2004). SimQuest discovery learning. Retrieved February 20, 2016 from http://www.simquest.nl

du Boulay, B., & Luckin, R. (2001). Modelling human teaching tactics and strategies for tutoring systems. *International Journal of Artificial Intelligence in Education*, *12*(3), 235-256.

Festinger L. 1957. *A theory of cognitive dissonance*. Stanford, CA: Stanford University Press.

Geier, R., Blumenfeld, P. C., Marx, R. W., Krajcik, J. S., Fishman, B., Soloway, E., & Clay-Chambers, J. (2008). Standardized test outcomes for students engaged in inquiry-based science curricula in the context of urban reform. *Journal of Research in Science Teaching*, *45*(8), 922-939.

Gordon, S. E., & Gill, R. T. (1997). Cognitive task analysis. In C.E. Zsambok, & G.A. Klein (Eds.), *Naturalistic decision making* (pp.131-140). Mahwah, NJ: Erlbaum.

Graesser, A. C., Person, N., Harter, D., & Tutoring Research Group. (2000). Teaching tactics in AutoTutor. In *Modelling human teaching tactics and strategies: Workshop W1 at ITS'2000, Montreal.*

Ha, S., & Sears, D. A. (2012). Using innovation with contrasting cases to scaffold collaborative learning and transfer. In J. van Aalst, K. Thompson, M. J. Jacobson, & P. Reimann (Eds.) *The Future of Learning: Proceedings of the 10th International Conference of the Learning Sciences (ICLS 2012) Conference Proceedings*, Vol. 1 (pp. 307-314).

Kapur, M. (2010). Productive failure in mathematical problem solving. *Instructional Science*, *38*(6), 523-550.

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science*, *40*(4), 651-672.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, *12*(1), 1-48.

Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review*, *19*(3), 239-264.

Lepper, M. R., Woolverton, M., Mumme, D. L., & Gurtner, J. (1993). Motivational techniques of expert human tutors: Lessons for the design of computer-based tutors. In S. P. Lajoie & S. J. Derry (Eds.), *Computers as cognitive tools: Technology in education*. Hillsdale, NJ: Erlbaum.

Linn, M. C. (1995). Designing computer learning environments for engineering and computer science: The scaffolded knowledge integration framework. *Journal of Science Education and Technology*, *4*(2), 103-126.

Linn, M. C., Clark, D., & Slotta, J. D. (2003). WISE design for knowledge integration. *Science Education*, *87*(4), 517-538.

Loibl, K., & Rummel, N. (2014). The impact of guidance during problem-solving prior to instruction on students' inventions and learning outcomes. *Instructional Science*, *42*(3), 305-326.

Mavrikis, M., & Gutierrez-Santos, S. (2010). Not all wizards are from Oz: Iterative design of intelligent learning environments by communication capacity tapering. *Computers & Education*, *54*(3), 641-651.

Mavrikis, M., Gutierrez-Santos, S., Geraniou, E., & Noss, R. (2013). Design requirements, student perception indicators and validation metrics for intelligent exploratory learning environments. *Personal and ubiquitous computing*, *17*(8), 1605-1620.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *American Psychologist*, *59*(1), 14-19.

Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences, 13(3), 273-304*.

Roll, I., Aleven, V., & Koedinger, K. R. (2009). Helping students know 'further'—increasing the flexibility of students' knowledge using symbolic invention tasks. In N. A. Taatgen & H. van Rijn (Eds.), *Proceedings of the 31st annual conference of the cognitive science society* (pp. 1169-1174). Austin, TX: Cognitive Science Society.

Roll, I., Aleven, V., & Koedinger, K. R. (2010). The invention lab: Using a hybrid of model tracing and constraint-based modeling to offer intelligent support in inquiry environments. In *Intelligent Tutoring Systems* (pp. 115-124). Springer Berlin Heidelberg.

Roll, I., Aleven, V., & Koedinger, K. R. (2011). Outcomes and mechanisms of transfer in invention activities. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd annual conference of the cognitive science society* (pp. 2824-2829). Austin: Cognitive Science Society.

Sandoval, W. A., & Reiser, B. J. (1997, March). *Evolving explanations in high school biology.* Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction*, *16*, 475–522.

Schwartz, D. L., Bransford, J. D., & Sears, D. (2005). Efficiency and innovation in transfer. In J. Mestre (Ed.), *Transfer of learning from a modern multidisciplinary perspective* (pp. 1–51). Greenwich, CT: Information

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.

Schwartz, D. L., & Martin, T. (2004). Inventing to prepare for future learning: The hidden efficiency of encouraging original student production in statistics instruction. *Cognition and Instruction*, *22*(2), 129-184.

Shemwell, J. T., Chase, C. C., & Schwartz, D. L. (2015). Seeking the general explanation: A test of inductive activities for learning and transfer. *Journal of Research in Science Teaching*, *52*(1), 58-83.

Stevens, A. L., & Collins, A. (1977). The goal structure of a Socratic tutor. In *Proceedings of the National ACM Conference* (pp. 256-263). Association of Computing Machinery.

van Joolingen, W. R., & de Jong, T. (1991). Supporting hypothesis generation by learners exploring an interactive computer simulation. *Instructional Science*, *20*(5-6), 389-404.

van Joolingen, W. R., & de Jong, T. (2003). Simquest. In *Authoring tools for advanced technology learning environments* (pp. 1-31). Springer Netherlands.

van Joolingen, W. R., de Jong, T., & Dimitrakopoulou, A. (2007). Issues in computer supported inquiry learning in science. *Journal of Computer Assisted Learning*, *23*(2), 111-119.

van Joolingen, W. R., de Jong, T., Lazonder, A. W., Savelsbergh, E. R., & Manlove, S. (2005). Co-Lab: Research and development of an online learning environment for collaborative scientific discovery learning. *Computers in Human Behavior, 21(4), 671–688*.

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal Artificial Intelligence in Education, 16*(3), 227–265.

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction, 21*(3), 209-249.

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89-100