Instr Sci (2012) 40:691–710
DOI 10.1007/s11251-012-9208-7

Evaluating metacognitive scaffolding in Guided Invention Activities

Ido Roll · Natasha G. Holmes · James Day · Doug Bonn

Received: 18 March 2011 / Accepted: 17 February 2012 / Published online: 21 March 2012

Abstract Invention and Productive Failure activities ask students to generate methods that capture the important properties of some given data (e.g., uncertainty) before being taught the expert solution. Invention and Productive Failure activities are a class of scientific inquiry activities in that students create, implement, and evaluate mathematical models based on data. Yet, lacking sufficient inquiry skills, students often do not actualize the full potential of these activities. We identified key invention strategies in which students often fail to engage: exploratory analysis, peer interaction, self-explanation, and evaluation. A classroom study with 134 students evaluated the effect of supporting these skills on the quality and outcomes of the invention process. Students in the Unguided Invention condition received conventional Invention Activities; students in the Guided Invention condition received complementary metacognitive scaffolding. Students were asked to invent methods for calculating uncertainties in best-fitting lines. Guided Invention students invented methods that included more conceptual features and ranked the given datasets more accurately, although the quality of their mathematical expressions was not improved. At the process level, Guided Invention students revised their methods more frequently and had more and better instances of unprompted self-explanations even on components of the activity that were not supported by the metacognitive scaffolding. Classroom observations are used to demonstrate the effect of the scaffolding on students' learning behaviours. These results suggest that process guidance in the form of metacognitive scaffolding augments the inherent benefits of Invention Activities and can lead to gains at both domain and inquiry levels.

I. Roll
Department of Educational and Counselling Psychology, and Special Education,
University of British Columbia, Vancouver, BC, Canada

I. Roll · N. G. Holmes · J. Day · D. Bonn
Department of Physics and Astronomy, University of British Columbia, Vancouver, BC, Canada

I. Roll (✉)
6224 Agricultural Road, Vancouver, BC V6T 1Z1, Canada
e-mail: ido@phas.ubc.ca

Keywords Invention Activities · Productive Failure · Metacognitive scaffolding · Inquiry learning · Guided discovery

Introduction

Constructivist pedagogies often include expository activities that prepare students for future instruction. For example, Productive Failure activities ask students to generate solutions to rich (and seemingly ill-structured) problems. Following the generation phase, a teacher-led consolidation phase helps students learn canonical solutions to the same problems (Kapur and Bielaczyc 2012). Another form of expository activities is Invention Activities. Invention Activities ask students to invent mathematical methods that capture meaningful properties of data prior to receiving instruction on canonical methods that achieve the same goal (Roll et al. 2009, 2011; Schwartz and Martin 2004). For example, students may be asked to invent a method for calculating statistical variability prior to learning about variance. Common to these approaches is the exposure of students to the problems prior to giving them tools to solve them. The early exposure helps students to appreciate the problem, to identify limitations of their prior knowledge, and in general, to prepare to learn from instruction (Roll et al. 2011; Schwartz and Bransford 1998). Also common to these approaches is the recognition of the value in errors; students' attempts to solve the problems are productive even though most students fail to invent valid solutions.

Expository activities that support knowledge construction prior to instruction have been shown to facilitate the acquisition of robust knowledge and to support transfer in a variety of contexts (Kapur and Bielaczyc 2011; Roll et al. 2011, Needham and Begg 1991; Schwartz and Martin 2004; Terwel et al. 2009). However, learning from constructivist activities is often less efficient than learning from tell-and-practice activities (Alfieri et al. 2011; Kirschner et al. 2006; Klahr and Nigam 2004; Tobias and Duffy 2009). In fact, research on a variety of constructivist activities suggests that students lack sufficient inquiry- and domain-level knowledge, and thus typically fail to use available resources appropriately (de Jong and van Joolingen 1998; Mulder et al. 2009). Thus, finding effective support for the generation process is of interest.

One potential form of support for students during expository activities is domain-level guidance. For example, students may be given suggestions and feedback regarding the outcomes of their invention process. This support is likely to help students succeed in the expository activity itself. However, succeeding in the task does not necessarily result in better learning. In fact, Kapur (2010) found that adding domain-level support increases the success of the generation process in Productive Failure activities, yet it hinders learning. Domain-level support may short-circuit important cognitive processes that are key to acquiring a deeper understanding of the domain and its features (Tobias and Duffy 2009).

In contrast with adding domain-level support, domain-independent support can help students engage with the invention process more productively, without reducing its constructivist characteristics. In this paper we evaluate the addition of metacognitive scaffolding, that is, domain-independent instructional prompts, to Invention Activities. Specifically, we study the effect of metacognitive scaffolding early in the invention activity on the *quality* and *outcomes* of students' invention process. We begin by describing Invention Activities and identifying key strategies that contribute to productive engagement with Invention Activities. Next, we present a classroom evaluation of the effect of metacognitive scaffolding for the target strategies. Last, we draw parallels between

Invention Activities and other forms of inquiry learning, and discuss implications for instructional design.

Invention Activities

Invention Activities have students invent methods that capture deep features of the domain using carefully designed data (Day et al. 2010; Roll et al. 2009, 2011; Schwartz and Martin 2004). For example, when given data sets and a common line of best fit, students are asked to invent a method for calculating the uncertainty in the slope (that is, to determine σ_m for the slope m of each data set, given that the line equation is $y = (m \pm \sigma_m)x$; see Fig. 1).

Invention Activities have students invent in small groups. Similar to Productive Failure activities (Kapur 2008, 2009), and unlike most constructivist activities, students are not expected to invent a correct method (Roll et al. 2009; Schwartz and Martin 2004). Instead, the invention process prepares students to learn better from subsequent instruction, thus augmenting, rather than replacing, show-and-practice instruction (Schwartz and Bransford 1998). Studies have repeatedly demonstrated that prefacing tell-and-practice instruction with Invention Activities results in better learning than tell-and-practice alone, controlling for overall time on task, especially on measures of transfer (Roll et al. 2009; Schwartz and Martin 2004). Invention Activities may also have domain-independent benefits, such as encouraging students to create multiple alternative solutions (Taylor et al. 2010). At the same time, Invention Activities do not seem to improve procedural ability (Roll et al. 2009, 2011; Taylor et al. 2010). Notably, Invention Activities contribute to learning even though students rarely succeed in inventing mathematically valid and general methods that capture the target properties (Schwartz and Martin 2004; Roll et al. 2009, 2011).

Contrasting cases in Invention Activities

Invention and Productive Failure activities include several forms of domain-independent support. Motivational support is given by creating a safe environment in which students' failed attempts are appreciated (Kapur 2009). Collaboration support may be offered by assigning roles to students (Westermann and Rummel 2012). In addition to domain-

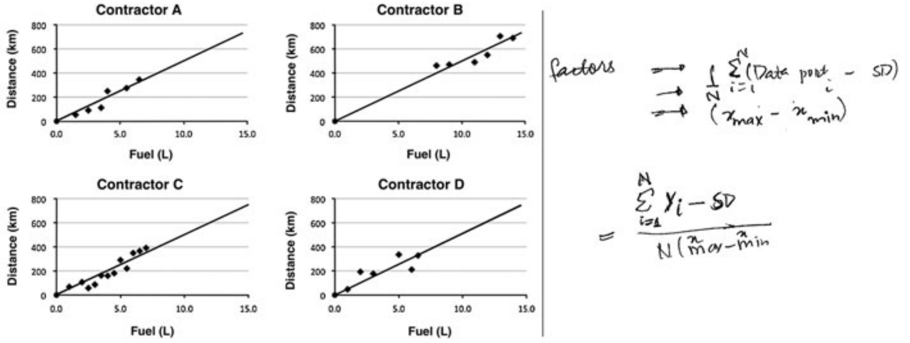


Fig. 1 A typical invention activity (*left*) and a student solution (*right*). Students are asked to invent a formula for calculating the uncertainty in the slope, σ_m. The data describe distance travelled (km) as a function of fuel consumed (L), and directs students' attention to the important features of the domain: range ($\sigma_{m,B} < \sigma_{m,A}$), sample size ($\sigma_{m,C} < \sigma_{m,A}$), and variability in the data ($\sigma_{m,D} > \sigma_{m,A}$)

independent support, significant domain-level support is provided in the form of the data given to students. Invention Activities include carefully designed examples that direct students' attention to the deep features of the target concept (Chase et al. 2010; Schwartz et al. 2007). Students compare and contrast the given examples to evaluate the relation between the emphasized features (e.g., sample size) and the target concepts (e.g., uncertainty). Hence these examples are referred to as *contrasting cases*. In reference to Fig. 1, and given that the line in this example must pass through the origin, students can notice the importance of *range* (data set A vs. B: data farther from the origin reduces the slope uncertainty); *sample size* (A vs. C: more data reduces the slope uncertainty); and *variability* in the data (A vs. D: greater variability increases the slope uncertainty).

Presenting contrasting cases simultaneously serves multiple purposes. First, the contrasts help students to notice structural variations that transcend superficial features (Schwartz et al. 2007; Star and Rittle-Johnson 2009). Second, students gain essential experiences by attempting to account for these features using symbolic methods (e.g., understanding that division by N accounts for sample size; Roll et al. 2011). Last, the contrasting cases offer opportunities for situational feedback (Mathan and Koedinger 2005; Nathan 1998). By applying the invented method to the given data sets students can evaluate whether their invented method provides a reasonable ranking of the contrasting cases, and, in the case of a failure, for which feature(s) the method fails to account. With regard to Fig. 1, values that are calculated by the invented method should reflect the fact that the uncertainty of data set A is visibly smaller than that of data set D and larger than data sets B and C.

Productive strategies in Invention Activities

One path to improving learning from Invention Activities may be by supporting students' invention process. Doing so requires understanding of how desired invention strategies, as applied by experts, differ from suboptimal invention strategies as applied by novices. To learn how experts and novices engage in the invention process we conducted empirical, descriptive and prescriptive, cognitive task analyses (Lovett 1998). The descriptive task analysis was done by recording and observing groups of students in first-year physics labs (Day et al. 2010) and introductory biology classes (Taylor et al. 2010) as they worked through Invention Activities. Several transcripts from these observations are presented in the "Results" section. The prescriptive task analysis was done by observing each of the authors of this paper thinking aloud when presented with Invention Activities within and outside their domain of expertise. A few of these activities were identical to the ones given to students while others were more advanced. Comparing the invention processes of experts and novices highlighted four inquiry strategies that are useful for Invention Activities:

Exploratory analysis

Learners should begin the invention process by analyzing the available data qualitatively. Exploratory analysis assists students in identifying the target features and in setting a baseline for evaluating their inventions. For example, comparing Graph A to Graph C in Fig. 1 highlights the effect of sample size on slope uncertainty. Schoenfeld (1992) found that a main difference between novices and experts is in the time they take to evaluate the task and the data prior to engaging in the solution process itself. Similarly, we often found that students did not engage in any planning or exploratory analysis phase, ignored the given data altogether, and applied their methods based solely on their prior knowledge. These students may transfer in familiar features, but they are not likely to gain appreciation for new features that are introduced by the data. For example, many students applied standard deviation (the familiar knowledge) to slope uncertainty (the target knowledge). However, transferring methods does not help students differentiate the target concepts from their prior knowledge. For instance, students who applied standard deviation did not realize the importance of rewarding data with more points (graph C vs. A in Fig. 1).

Peer interaction

Experts who engage in Invention Activities often discuss their reasoning with their peers. Learning is facilitated not only by receiving feedback on one's own ideas, but also by attending to the ideas of others (White and Frederiksen 1998). In addition, by describing their often intuitive and informal methods and insights, students identify gaps in their reasoning. Yet, even though productive collaboration can be fruitful, students were rarely observed interacting with peers from other groups, and interaction within the groups was often limited. For example, a few groups had only one active member, or group members within the same group worked individually.

Self-explanation

Students who self-explain their reasoning acquire a better understanding of the domain (Chi et al. 1994), even when explaining erroneous answers (Siegler 2002). The use of self-explanations to enhance understanding can be seen in the solution shown in Fig. 1. The students first identified the key features and then used these as goals to direct the design of their method. However, students were often observed making conjectures based on intuition, prior knowledge, or social conformity, without justifying or explaining these.

Evaluation of outcomes

Experts often engage in an evaluation of their methods by triangulating their model-generated values with their initial observations. Evaluating their methods helps learners realize the limitations of the methods and triggers an iterative "debugging" process in which students look for better methods that achieve their intended goals. In the context of Productive Failure activities, there was a positive and significant correlation between the number of different solution methods and learning gains (Kapur 2012; Kapur and Bielaczyc 2011). However, in our observations it was clear that students often stop developing their methods as soon as their invented methods provide numerical outputs, falling short of evaluating them. While reaching a single solution is often acceptable in the classroom context, the invention process depends on an iterative design, facilitated by situational feedback.

Students were often found to skip a combination of these strategies. For example, students who do not engage in exploratory analysis, also tend to make unexplained (and often unjustified) conjectures, since their reasoning tends to be ad-hoc, without attempting to relate to the given data. Later these students often do not evaluate their outcomes, since they have not established a baseline during the exploratory analysis.

Supporting students' use of inquiry strategies

The infrequent use of the identified strategies by novice students emphasizes the need for explicit, domain-independent support for the invention process (Azevedo and Jacobson 2008; de Jong 2006). Overall, it seems that students are in need for support mainly during the exploratory data analysis phase and prior to the generation of methods per se.

One successful form of support is metacognitive scaffolding. Metacognitive scaffolding guides students by asking questions about the task or suggesting relevant domain-independent strategies (Bulu and Pedersen 2010; Ge and Land 2003; Osman and Hannafin 1994; Roll et al. 2007). Metacognitive scaffolding is different from domain-level scaffolding in that it focuses on the stage of the task and not on specific content. Thus, metacognitive scaffolding is reusable across different problems with similar structure. Metacognitive scaffolding was found to improve students' strategy use whether given by a human tutor (Azevedo et al. 2004), an intelligent tutoring system (Koedinger et al. 2009; Roll et al. 2011), or embedded in the task itself (Bulu and Pedersen 2010). However, while metacognitive scaffolding was found to improve strategy use, it often did not improve learning outcomes at the domain level (Bulu and Pedersen 2010; Manlove et al. 2007; Roll et al. 2011).

Within the scope of the current study we evaluated the effect of the metacognitive scaffolding on students' invention behaviors and outcomes in an in vivo quasi-experiment. The first research question is: what are the effects of metacognitive scaffolding on the quality of the invented methods? We hypothesize that students who received metacognitive scaffolding would be more likely to attend to deep features, and thus would incorporate more deep features in their invented methods. At the same time, since Invention Activities are not designed to improve procedural fluency, we hypothesize that the metacognitive scaffolding would not improve the mathematical quality of the methods.

The second research question is: what are the effects of metacognitive scaffolding on students' use of unsupported strategies? For example, will the given scaffolding also encourage students to evaluate their outcomes more frequently? Given the short manipulation, we hypothesize that the effect of the metacognitive scaffolding would be local and would not alter students' behaviors on unsupported aspects of the invention activity.

Method

Participants

The quasi-experiment took place in four sections of a first-year undergraduate physics laboratory course at the University of British Columbia. Two sections with 58 students were assigned to the Unguided Invention condition (control), and two sections with 76 students were assigned to the Guided Invention condition (treatment). Section assignment tried to control for prior knowledge and time of day. The majority of the students were 17–20 years of age and in the first year of their undergraduate degrees. 59.9% of the participants were males. All students had previously learned techniques for estimating uncertainties in data (e.g., standard deviation), but had not learned about uncertainties in slopes. Students were familiar with Invention Activities, with previous activities focusing on graphical representations and simpler statistical concepts (e.g., linear least-squares fitting; Day et al. 2010).

Materials

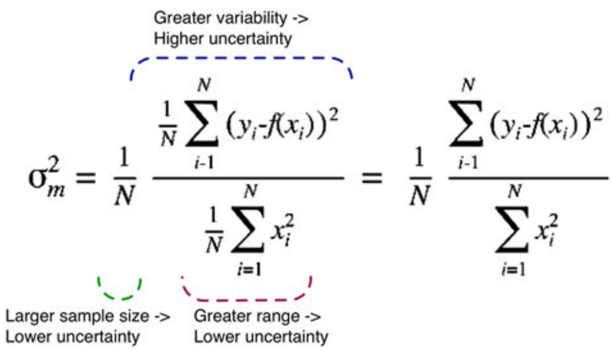
Students in both conditions were given an invention activity targeting the concept of uncertainty in slopes (see Fig. 1). The cover story described four contractors who measured the fuel efficiency of a single vehicle. Although the measured data varied for each contractor, all data sets were linear with a slope of 50 km/L, representing the fuel efficiency of the vehicle, and had a y-intercept of zero, since the vehicle would not be able to travel without any fuel. The goal of the activity was to invent a method that would determine the uncertainty in the slope from each of the data sets; that is, to find a method for calculating σ_{m}, given that the equation for distance (y) as a function of volume of fuel (x) is $y = (50 \pm \sigma_{\mathrm{m}})^{*}x$. The expert solution to this activity is presented in Fig. 2. The data sets were created so that specific pairwise comparisons highlighted the three features contained in the expert formula for calculating the uncertainty in a slope:

- *Variability in data*: the greater the distance between the data points and the line of best fit, the higher the uncertainty in the slope.
- *Range*: the farther out data points are along the x-axis, the smaller the uncertainty in the slope.
- *Sample size*: the more measurements collected, the smaller the uncertainty in the slope.

The Guided Invention condition differed from the Unguided Invention condition in that it included metacognitive scaffolding to address three of the target inquiry strategies:

- To engage students in exploratory analysis, students were instructed to qualitatively analyze the data by ranking the different contrasting cases. For three key pairwise comparisons (A–B, A–C, and A–D), students were asked: "whose data set does a better job of measuring the slope"? Additional prompts had students rank the contrasting cases in order of lowest-to-highest uncertainty in the slope.
- To support self-explanation, students were explicitly asked to explain their reasoning during the pairwise comparisons.
- To engage students in peer-interaction, students were asked to communicate their findings from the exploratory analysis phase, prior to designing their methods: "spend a few minutes comparing your answers to the other pairs at your table".

The fourth strategy, evaluation of outcomes, was assessed, but was not supported explicitly by the metacognitive scaffolding.



$$\sigma_m^2 = \frac{1}{N}\frac{\frac{1}{N}\sum_{i-1}^{N}(y_i - f(x_i))^2}{\frac{1}{N}\sum_{i=1}^{N}x_i^2} = \frac{1}{N}\frac{\sum_{i-1}^{N}(y_i - f(x_i))^2}{\sum_{i=1}^{N}x_i^2}$$

Fig. 2 The expert solution includes three components that correspond to the three target features: variability, range, and sample size

Table 1 Alignment between the inquiry strategies, metacognitive scaffolding, and assessment

	Inquiry strategy			
	Exploratory analysis	Self-explanation	Peer interaction	Evaluation
Treatment (Guided Invention condition only)	Engage in pairwise comparison; rank data sets	Explain pairwise comparisons	Discuss with peers	
Assessment				
Quantitative data		# and topic of high-level written comments		# of revised methods
Qualitative data	Examples from transcripts		Examples from transcripts	Examples from transcripts

The quality of students' invented methods was also assessed

The metacognitive scaffolding was given only during the exploratory analysis phase (see Table 1). Students' use of the strategies was evaluated during subsequent phases of the activity.

Design and procedure

Students in both conditions first received an introduction, delivered by the course instructor (the fourth author), which tied the goal of the invention activity to the course learning goals. Students were then given the contrasting cases, and those in the Guided Invention condition also received the designated metacognitive scaffolding. Students in both conditions were then asked to "write down a formula for calculating the uncertainty in the slope, σ_m" and to use their formula to calculate the uncertainty in the slope for each of the data sets provided. Students worked in randomly assigned pairs, though each student was required to hand in a worksheet. Students were given up to 1 h to work through the activity. Following the invention process, the instructor gave a short lecture describing the canonical formula for the uncertainty in the slope of a linear model with a y-intercept equal to zero (see Fig. 2). Students then practiced using this formula with empirical data they collected from a laboratory experiment.

Assessment plan

Students' invented methods were analyzed to evaluate the effect of the scaffolding on the invented methods and the invention process.

Quality of invented methods

A method that captures the deep features of the domain would accurately rank the relative uncertainties of the given contrasting cases ($\sigma_{m,B} < \sigma_{m,A}$; $\sigma_{m,C} < \sigma_{m,A}$; and $\sigma_{m,D} > \sigma_{m,A}$). Thus, the number of correct key pairwise rankings that are produced by a method is a good metric for measuring the conceptual validity of the method. Each method received a score of 0–3 based on the number of correct key pairwise comparisons it produced.

A second analysis evaluated how well students identified and appreciated each of the deep features of the target concept. A method that includes more of the target features is assumed to reflect a better understanding of the deep structure of the domain. Thus, each method received a score of 0–3 based on the number of deep features it included (range, spread, and sample size).

The third analysis of the invented methods was more technical, and evaluated the quality of the mathematical expressions of each method. For example, to get full credit when including range, a method should include some representation of x in the denominator, use all of the data, make all values positive by squaring them, and average the squared values. The student solution in Fig. 1 included x in its denominator, but it used only two points (x_{min} and x_{max}) and it neither squared nor averaged the outcome. A similar analysis was done for each of the features, totaling in 16 required technical features. Each method received a score based on the proportion of technical requirements that were valid (0–1).

Use of inquiry strategies

While students in both conditions were instructed to "write down a formula" and no self-explanation prompts were given during the planning phase, most students spontaneously included some form of written comments with their invented solutions. We considered any text that was not part of the method to be a comment. Students' comments were categorized as either *superficial comments*, which merely explained what the formula would do (e.g., "then we took a square root …") or *high-level* comments. High-level comments could include plans (e.g., "farther from the origin [should give] more weight"), justifications (e.g., "… and therefore a larger data set will have a smaller uncertainty"), and evaluations (e.g., "we went with [a second formula] because the first model gave what we thought was an unrealistic uncertainty for contractor D").

Each method was analyzed with regard to (i) whether it included any comments (the solution in Fig. 1 includes a comment); (ii) the number of high-level comments that were associated with the method (the solution in Fig. 1 has three high-level comments); and (iii) whether students chose to comment on the target features or on properties that were not targeted by the activity, namely, magnitude and units of the calculated values (all comments in Fig. 1 focus on the target features).

A few groups developed more than one method, often crossing out earlier versions in favor of more updated ones. As noted earlier, multiple solutions are often evidence of engaging in an evaluation process, and correlate with pre-to-post learning gains. Thus, we counted the number of methods that were developed by each group.

Last, selected transcripts from the classroom are used to demonstrate the effect of the metacognitive scaffolding. These chosen transcripts were typical to the groups that were recorded. While the detailed transcripts are only examples, they offer another perspective in understanding the effect of metacognitive scaffolding. Analysis of students' use of the target strategies is summarizes in Table 1.

Students' background knowledge with respect to basic statistics and data handling ability was assessed using the Concise Data Processing Assessment (Day and Bonn 2011). The assessment was administered to students several weeks after the study, and did not test knowledge of uncertainty in slopes.

The effect of condition was evaluated using ANCOVAs with condition as a factor and background knowledge as a covariate. Where the dependant variable was proportion of students, we used binomial probabilities test. A p-level of 0.05 was used in all analyses.

Results

There were no significant differences between conditions with regard to background knowledge: guided invention: $M = 0.42$, $SD = 0.21$; Unguided Invention: $M = 0.38$, $SD = 0.22$; $t(130) = 1.05$, $p = 0.29$.

Quality of invented methods

The first analysis evaluated the number of correct key pairwise comparisons based on the values that were computed by each method. The activity included three key pairwise comparisons (see Fig. 1), and thus each method received a score of 0–3. Methods invented by the Guided Invention students ranked the uncertainties significantly better than the methods invented by the Unguided Invention students (2.1 vs. 1.8, respectively), $F(2,131) = 6.1$, $p = 0.01$ (see Table 2).

Students' encoding of the domain is also reflected in the number of deep features that were incorporated into each invented method. An invented method received a score of 0–3 based on the number of deep features included in the method (variability, range, and sample size). Students in the Guided Invention condition incorporated on average 2.05 features, compared with 1.83 features by students in the Unguided Invention condition. This effect is marginally significant, $F(2,131) = 3.48$, $p = 0.06$. A separate analysis for each feature found that Guided Invention students were three times more likely to include sample size in their methods (guided invention: $M = 0.42$; Unguided Invention: $M = 0.14$), $\chi^2(1, N = 134) = 8.8$, $p = 0.003$. There was no significant difference in the other features (variability: M = 0.99 and 0.98 respectively; range: M = 0.65 and 0.71 respectively).

The third analysis evaluated the technical quality of the mathematical expressions students invented. Condition was not a significant factor in determining the technical quality of the invented methods: guided invention: $M = 0.56$, Unguided Invention:

Table 2 Quality of invented methods and unprompted application of inquiry strategies: M (SD)

	Guided Invention	Unguided Invention
Quality of invented methods		
# of correct rankings of contrasting cases (0–3)	2.12 (0.68)	1.79 (0.78)**
# of included features (0–3)	2.05 (0.71)	1.83 (0.60)†
Quality of mathematical expressions (0–1)	0.56 (0.19)	0.51 (0.17)
Unprompted application of inquiry strategies		
Self explanation		
# of comments per student	1.03 (0.82)	1.05 (0.69)
# of high-level comments per student	0.96 (1.2)	0.55 (0.90)*
# of high-level comments that focus on target features per student	0.91 (1.1)	0.34 (0.66)**
Evaluation		
% of students with multiple alternative methods (%)	13	3*

† $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

$M = 0.51$, $F(2,131) = 1.99$; $p = 0.16$. None of the students invented a method that was applicable to all possible data sets.

Students' use of inquiry strategies

Exploratory data analysis

The metacognitive scaffolding given in the Guided Invention condition instructed students to engage in exploratory analysis by asking them to rank the key pairwise comparisons. The following transcript of Guided Invention students demonstrates how engaging in exploratory analysis helped students to notice and explain the role of range:

Excerpt 1
s1: So we need to discuss these questions I guess [Pointing to the question that asks to rank the uncertainties of sets A and B] … Okay so A or B?
s2: I think…
s1: [Pointing to Graph A] These ones all have [their data] before 8 [on the x-axis], and these [pointing to Graph B] are all after 8 [on the x-axis].
s2: I would say, the…
s1: There's more…
s2: The bigger the distance [from the origin, along the x-axis], the less the uncertainty.
s1: Really? The bigger the distance?
s2: Yeah. Because if you want to produce a slope, with this [points to Graph A], this is the maximum uncertainty here, this dot [the furthest from the line in Graph A]. … and the angle between this line [between this data point and the origin] and the line of best fit is going to be smaller than if you take this dot [pointing to the data point furthest from the line in Graph B].

The excerpt demonstrates how students were guided by the metacognitive scaffolding to engage in exploratory analysis, and how this process resulted in noticing key features of the domain. The two students in this example had not realized the importance of range prior to comparing data sets A and B, but the perceptual comparison of the two data sets directed their attention to range. S1 noticed the feature, and s2 explained its importance using qualitative reasoning: the farther a point is from the origin, the less the line defined by it and the origin will deviate from the best-fit line.

Peer interaction

The scaffolding asked students to discuss their ranking of data sets with peers from other groups. The two groups in the excerpt below debate which feature is more important: range (as emphasized by comparing Graph A vs. B) or sample size (Graph A vs. C):

Excerpt 2
s1 (*Group 1*): [Turning to another group] Okay, so we have to compare our answers.
s2 (*Group 1*): What do you say?
s3 (*Group 2*): C–B–A–D
s1 (*Group 1*): OK, we said, B–C–A–D.
s2 (*Group 1*): C–B–A–D, and why?
s3 (*Group 2*): C has more data points…
s2 (Group 1): Yeah, but it's all in the region close to the origin. But, yeah, I mean…

s1 (*Group 1*): Well, we thought that since B has more data points farther out, that that would mean, well; basically, what we said was the ratio of the vertical distance over the actual distance traveled was smaller in here [in B], than in C. But I don't know if that's more important than the number of data points.
s3 (*Group 2*): I'm starting to think B now [corresponds to range], because, I just read there's no intercept.

As shown in the discussion above, neither student in Group 2 realized the importance of range until explained by s1 (Group 1). Students also engaged in the important discussion of the trade-off between different features, thus developing an integrated and coherent understanding of the overall structure of the domain.

Self-explanation

Most students in both conditions spontaneously included some form of written comments with their invented solutions, even though no self-explanation prompts were given at this stage. Overall 90% of the students wrote comments, with no effect for condition. The average number of comments per student was 1.03 in the Guided Invention condition, and 1.05 in the Unguided Invention condition; $F(2,131) = 0.04$, $p = 0.84$. However, not all comments are equally important. When looking at the level of students' comments, condition had a significant effect on the average number of *high-level* comments. Guided Invention students made almost twice as many high-level comments per student, $M = 0.96$, compared with the Unguided Invention students, $M = 0.55$; $F(2,131) = 4.0$, $p = 0.05$.

High-level comments were made with respect to five features of the invented methods: The three target features (variability, range, and sample size), and two other features that were not the focus of the activity (the magnitude of the calculated values and their units). Condition had a significant effect on the focus of students' comments: Students in the Guided Invention condition made on average 0.91 comments per student on target features, contrasted with 0.34 comments per student in the Unguided Invention condition; $F(2,131) = 10.7$; $p = 0.001$, see Fig. 3.

Evaluation

Students' tendency to evaluate their methods was assessed by counting the number of methods that each student wrote down. The presence of multiple methods suggests that students evaluated previous methods and found them unsatisfactory. Overall, 13% of the students in the Guided Invention condition presented multiple solutions, compared to 3% of students in the Unguided Invention Condition. The effect of condition on number of methods is significant, $F(2,131) = 5.6$, $p = 0.02$.

The class observations demonstrate how the metacognitive scaffolding encouraged students to evaluate their methods. The following excerpt is taken from students who were implementing their method using the given data. The students in this example made an observation during their exploratory analysis that the uncertainty in data set C is lower than that in data set A, since C has more data points. Following their invention attempts they put their method to the test:

Excerpt 3
s1: Now we see if our reasoning is of any worth.
s2: Oh yeah, true. Okay.

s*1*: This [uncertainty for C] is bigger [than uncertainty for A], we screwed up.
s2: Is that good? This is C…
s1: Actually this is not good, because if this [uncertainty for C] is bigger than this one [uncertainty for A]…
s2: Oh yeah, and also we didn't take into account the number of data points [for C].
s1: Why? We divided by 12
s2: So, yeah, our uncertainty doesn't take into account how many data points there are.
s1: No it does! The more data points…
s2: But we added them up and divided by the numbers: so, we took the average. So, the number of data points doesn't matter, you know what I mean?

Excerpt 3 shows how the evaluation process helped students to identify a limitation of their method. The evaluation further helped the students to learn about the applicability of different mathematical manipulations. Specifically, the students experienced that division by N may not reward data sets with more points, and might simply compensate for summation of a different quantity of addends.

Discussion

Supporting students during the invention and generation process

Analysis of students' invented methods shows that students in the Guided Invention condition invented methods that captured the ranking of the given data more accurately and incorporated more target features in their methods. While the metacognitive scaffolding improved the conceptual quality of the invented methods, it had no effect on the technical quality of the mathematical expressions. Excerpts from student interactions demonstrate how the metacognitive scaffolding achieves these gains. It seems that the scaffolding augments the strengths of Invention Activities (better conceptual understanding), without contributing to other aspects of students' invented methods (technical validity). These results support our hypotheses with regard to the invention outcomes.

Contrary to our second hypothesis, the metacognitive scaffolding also had a positive effect on students' use of unprompted inquiry strategies. Guided Invention students made more high-level comments per student, in which they were more likely to refer to the target features of the activity. Students in the Guided Invention condition were also more likely to evaluate their methods, and subsequently, to develop more solutions. Given that number of methods is often correlated with quality of learning (e.g., Kapur 2012; Wiedmann et al. 2012), increasing the number of alternative solutions is an important step towards supporting more expert-like behaviors. The increased number of methods echoes a similar effect for Invention Activities in Biology classrooms (Taylor et al. 2010). It is important to emphasize that students were neither prompted to self-explain their methods, nor to evaluate them and develop multiple solutions. The fact that scaffolding particular strategies led to improvement on related strategies (or on the same strategies in different phases of the activity) suggests that the metacognitive scaffolding put students in a more "metacognitive mindset", thus achieving an effect that carries over to other aspects of the task.

The effect of metacognitive scaffolding on learning outcomes

The effect of the metacognitive scaffolding on learning from the subsequent instruction and practice (or consolidation) phases is not clear. Invention attempts were previously found to improve learning from subsequent instruction, even though students failed to invent valid methods. Will improving students' ability to invent hinder their ability to learn from the subsequent instruction? Doubtfully. The Invention and Productive Failure literatures do not argue that students *should* be failing. In fact, the discovery learning literature shows that succeeding in generating solutions leads to robust learning (McDaniel and Schlager 1990). In other words, the Invention and Productive Failure literatures suggest that some failures are productive—not that succeeding is counterproductive. Students learn from invention and Productive Failure activities by gaining experiences that help them differentiate prior knowledge and encode the subsequent instruction (Schwartz et al. 2007). For example, Invention Activities help students set concrete requirements from the canonical solutions (Roll et al. 2011). The metacognitive scaffolding helps students to notice more of the target features, and thus set better goals for the canonical method. We therefore hypothesize that metacognitive scaffolding will lead to better learning from the subsequent instruction. Yet, this remains to be evaluated.

Inventing as a process of inquiry

The invention process is similar to scientific inquiry in that data is used to make inferences, construct models, and validate them by comparing outputs to observations (Godfrey-Smith 2003). As shown above, this process requires students to apply several inquiry-level, domain-independent skills. Data collected in the study allows us offer a better operational definition of the invention process. Table 3 compares the main stages of the invention process with various frameworks of scientific inquiry and problem solving.

The first stage of the invention process is *task definition*. Specifically, students should realize that their goal is to invent mathematical models that capture a specific property of data (e.g., uncertainty) and works across the given data. These learning challenges are designed to resemble the challenges that scientists face (de Jong 2006; Godfrey-Smith 2003).

The second stage is *analysis*. Students should analyze the available data and identify patterns that correspond to the structure of the target concept. This analysis can be done by comparing data sets (as is done in Invention Activities) or by observing a phenomenon (Godfrey-Smith 2003). De Jong (2006) refers to this stage as orientation, since the learner familiarizes herself with the domain and its features. The outcomes of this phase are concrete observations that highlight key features of the target concept (e.g., a larger sample size reduces statistical uncertainty).

Table 3 Stages in the invention process compared with inquiry learning, problem solving, and the scientific method

Invention Activities	Inquiry learning (de Jong 2006)	Self regulation during mathematical problem solving (Schoenfeld 1987)	The scientific method (Godfrey-Smith 2003)
Task definition		Read	Question formation
Analysis	Orientation	Analyze	Analysis
Plan and design	Hypothesis generation	Plan	Explanation
Implementation	Experimentation	Implement	
Evaluation	Conclusion and evaluation	Verify	Prediction and evaluation

The third stage is *plan and design*, during which students create their methods by hypothesizing which combination of mathematical tools captures the identified features and explains their initial observations. For example, students may realize that division by N is a mathematical tool that reduces variability for a larger sample size.

The fourth stage is *implementation* of the designed plans. In this step students apply their methods to the available data (in Invention Activities) or execute their plans (in discovery learning and problem solving).

The last stage is *evaluation*, in which students evaluate their methods by comparing their outcomes (step 4) to their initial observations (step 2).

While these five stages have been ordered sequentially, they are iterative and intertwined. For instance, students who realize that their method does not align with their initial observations may adapt their method or go back to the data analysis phase in order to reassess their goals.

The role of exploratory data analysis

The metacognitive scaffolding that was given to students focused on the exploratory analysis phase. Students were instructed to make pairwise comparisons, explain their reasoning, and discuss their tentative conclusions with their peers—all of which happened prior to attempting to invent their methods.

The effect of the metacognitive scaffolding on students' invention behaviors and outcomes establishes the importance of the exploratory analysis to the symbolic invention process. This result echoes earlier findings about the importance of planning in math problem solving (Schoenfeld 1987). Schwartz et al. (2007) suggest that the invention process helps students acquire relevant experiences that prepare them for future learning. To support learning from subsequent instruction, not only should students notice the deep features, but also they should appreciate their relevance to the target concept. As shown in Fig. 3, comments written by the Guided Invention students were 3 times more likely to focus on the target features, suggesting that the metacognitive scaffolding helped students to develop an appreciation of the target concepts.

Contrasting two forms of reasoning

It seems that a productive invention process includes several forms of contrasts. The first form of contrasts is between the given data sets. These contrasts help students to notice the deep features of the domain. A second level of contrasts is between the different student-invented methods. Contrasting methods allows students to evaluate how data behaves under different symbolic manipulations and which procedures can best account for the identified features, much like contrasting alternative solution approaches (Star and Rittle-Johnson 2009). Yet, perhaps the most interesting form of contrast is the process of coordination between two forms of reasoning: qualitative exploratory analysis and quantitative design of symbolic methods.

The invention process includes two forms of reasoning with data: a qualitative, exploratory data-analysis phase and a quantitative, symbolic design and evaluation phase. The benefits of the symbolic component of the invention process have been demonstrated elsewhere. For example, previous studies showed that engaging in the exploratory analysis phase without the symbolic invention phase is less effective than engaging in both phases (Kapur and Bielaczyc 2011; Roll et al. 2009, 2011). In a different domain, symbolic reasoning with contrasting cases was found to lead to better understanding of the domain,

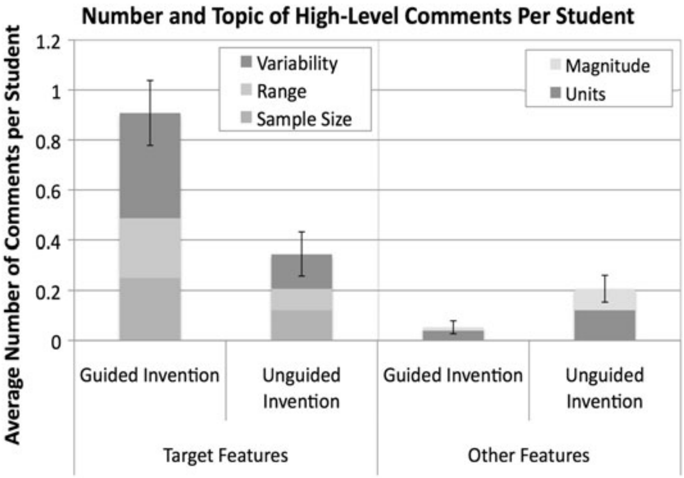


Fig. 3 Number and topic of high-level comments per commenting student as a function of condition. Guided Invention students made more high-level comments, and their comments focused more on target features of the domain

compared with the verbal analysis of the same contrasting cases (Schwartz et al. 2005). These studies establish the importance of symbolic reasoning to students' qualitative understanding of the domain. On more than one occasion we noticed students who used their invented methods to go back and reflect upon (and update) their qualitative analysis.

The current study, however, focuses on the benefits of *qualitative* analysis (in the form of exploratory analysis) to quantitative reasoning. The qualitative analysis helps students to notice the deep features of the domain and offers opportunities to evaluate the invented methods. It seems that the benefits of the invention process are not due to any single reasoning process. Rather, it is the interaction between these two complementary forms of reasoning that helps students realize what features are critical in the domain and what symbolic manipulations can account for them. Learning from Invention Activities happens not only because students attend to deep features during the qualitative analysis and experiment with mathematical formulas during the quantitative invention process. Rather, we suggest that learning happens mainly when one process reciprocally informs the other (see Fig. 4).

Summary

Invention Activities, like Productive Failure activities, ask students to invent novel solutions to unfamiliar problems prior to receiving instruction. We identify inquiry strategies that help students make sense of data and draw parallels between the invention process and inquiry learning. We further describe an evaluation of the effect of metacognitive scaffolding on the invention process and its outcomes. Results show that the metacognitive scaffolding increased the quality of the invented methods, the number and quality of unprompted self-explanations, and the frequency of multiple solutions. From an instructional design perspective, it seems that metacognitive scaffolding can be a powerful tool in helping students to improve their inquiry behaviors and to learn from constructivist

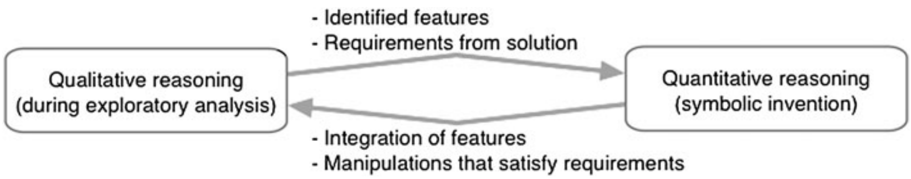


Fig. 4 Co-evolution of qualitative and quantitative reasoning. The products of each process feed into the other and support development of relevant mental models

activities. Within the context of this study, prompting students to use inquiry strategies during the invention process put them in a more metacognitive mindset, which benefited the later components of the activity.

The Invention and Productive Failure literatures emphasize the importance of exposing students to the challenges of the domains prior to giving them tools to overcome these challenges. This study emphasizes the benefits of contrasting cases and suggests that instructors should encourage students to reason using multiple approaches: a qualitative, often perceptual, reasoning, and symbolic reasoning using mathematical notations. The iterative comparisons between data sets, solution approaches, and forms of reasoning allow students to identify key information in the given data, extract feedback regarding their inquiry behavior, and inform their understanding of the new domain.

Unlike previous studies that evaluated learning from Invention Activities, the current study analyzed the actual methods that students invented. This analysis demonstrates how students who attend to the contrasting cases identify key features and construct an integrative symbolic (and likely also mental) model of the domain. These results suggest that instructors should consider using Invention Activities in conceptual domains, where students lack prior knowledge and understanding of the structure.

The study has several limitations. First and foremost, while it presents a detailed analysis of the invention process and outcomes, it does not evaluate students' learning, or the effect of the metacognitive scaffolding on the subsequent instruction and practice phases. It is likely that the quality of students' learning correlates with the extent to which they notice and incorporate deep features in their invented methods, however, this assumption is not tested in the scope of this study. A second limitation has to do with the long-term effects of metacognitive scaffolding. It may be that students learn to rely on the scaffolding and thus become less independent learners. The metacognitive scaffolding may assist performance on the activity being supported, but hurt ability to deal with similar, unsupported activities in the future. Alternatively, the metacognitive scaffolding may function like a metacognitive worked-out example, thus teaching students how to approach the invention process. A hint in that direction can be found in the fact that students improved their self-explanation behaviors even without relevant scaffolding. Bulu and Pedersen (2010) and Roll et al. (2011) found that the benefits of metacognitive scaffolding could be retained even after support is removed. The third limitation is the indirect inferences on inquiry behaviors from properties of the invention outcomes. While the classroom observations give a window into students' inquiry behaviors, a more detailed analysis of the transcripts from the observations is desired. Last, the scope of the paper is limited to university level data analysis skills.

This study makes several contributions. First, it offers an operational definition of the invention process and situates Invention Activities within other frameworks of inquiry learning. Second, it demonstrates the fashion in which attending to contrasting cases helps students to notice and appreciate the deep features of the target domain. Third, it demonstrates the importance of employing a qualitative reasoning process alongside the quantitative one. Last, it puts forward a set of metacognitive scaffoldings that improve the quality of the invented methods and the corresponding reasoning processes. It seems that while Productive Failure and Invention Activities demonstrate the benefits of withholding information at the domain level, supporting students at the metacognitive level may, in fact, increase the productivity of student's failures.

Acknowledgments This work was supported by the Pittsburgh Science of Learning Center, which is supported by the National Science Foundation (#SBE-0836012), and by the University of British Columbia through the Carl Wieman Science Education Initiative.

References

Alfieri, L., Brooks, P. J., Aldrich, N. J., & Tenenbaum, H. R. (2011). Does discovery-based instruction enhance learning? *Journal of Educational Psychology, 103*(1), 1–18.

Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology, 29*, 344–370.

Azevedo, R., & Jacobson, M. J. (2008). Advances in scaffolding learning with hypertext and hypermedia: A summary and critical analysis. *Educational Technology Research and Development, 56*, 93–100.

Bulu, S., & Pedersen, S. (2010). Scaffolding middle school students' content knowledge and ill-structured problem solving in a problem-based hypermedia learning environment. *Educational Technology Research and Development, 58*(5), 507–529.

Chase, C. C., Shemwell, J. T., & Schwartz, D. L. (2010). Explaining across contrasting cases for deep understanding in science: An example using interactive simulations. In S. Goldman & J. Pellegrino (Eds.), *Proceedings of the 9th international conference of the learning sciences* (pp. 153–160). Chicago, IL.

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

Day, J., & Bonn, D. (2011). Development of the concise data processing assessment. *Physical Review Special Topics Physics Education Research, 7*(1), 010114.

Day, J., Nakahara, H., & Bonn, D. (2010). Teaching standard deviation by building from student invention. *The Physics Teacher, 48*, 546–548.

de Jong, T. (2006). Scaffolds for scientific discovery learning. In J. Elen, R. E. Clark, & J. Lowyck (Eds.), *Handling complexity in learning environments: Theory and research* (pp. 107–128). Howard House: Emerald Group Publishing.

de Jong, T., & van Joolingen, W. R. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research, 68*, 179–201.

Ge, X., & Land, S. M. (2003). Scaffolding students' problem-solving processes in an ill-structured task using question prompts and peer interactions. *Educational Technology Research and Development, 51*(1), 21–38.

Godfrey-Smith, P. (2003). *Theory and reality: An introduction to the philosophy of science*. Chicago: University of Chicago Press.

Kapur, M. (2008). Productive failure. *Cognition and Instruction, 26*(3), 379–424.

Kapur, M. (2009). Productive failure in mathematical problem solving. *Instructional Science, 38*(6), 523–550.

Kapur, M. (2010). A further study of productive failure in mathematical problem solving: Unpacking the design components. *Instructional Science, 39*(4), 561–579. doi:10.1007/s11251-010-9144-3.

Kapur, M. (2012). Productive failure in learning the concept of variance. *Instructional Science* (this issue).

Kapur, M., & Bielaczyc, K. (2011). Classroom-based experiments in productive failure. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (pp. 2812–2817). Austin: Cognitive Science Society.

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *The Journal of the Learning Sciences, 21*(1), 45–83.

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.

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.

Koedinger, K. R., Aleven, V., Roll, I., & Baker, R. S. J.d. (2009). In vivo experiments on whether supporting metacognition in intelligent tutoring systems yields robust learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 383–412). New York: Routledge.

Lovett, M. (1998). Cognitive task analysis in service of intelligent tutoring system design: A case study in statistics. In B. P. Goettl, et al. (Eds.), *Intelligent Tutoring Systems Lecture Notes in Computer Science* (pp. 234–243). Berlin: Springer.

Manlove, S., Lazonder, A. W., & de Jong, T. (2007). Software scaffolds to promote regulation during scientific inquiry learning. *Metacognition and Learning, 2*, 141–155.

Mathan, S. A., & Koedinger, K. R. (2005). Fostering the intelligent novice: Learning from errors with metacognitive tutoring. *Educational Psychologist, 40*(4), 257–265.

McDaniel, M. A., & Schlager, M. S. (1990). Discovery learning and transfer of problem-solving skills. *Cognition and Instruction, 7*(2), 129–159.

Mulder, Y. G., Lazonder, A. W., & de Jong, T. (2009). Finding out how they find it out: An empirical analysis of inquiry learners' need for support. *International Journal of Science Education*, 1–21.

Nathan, M. J. (1998). Knowledge and situational feedback in a learning environment for algebra story problem solving. *Interactive Learning Environments, 5*(1), 135–159.

Needham, D. R., & Begg, I. M. (1991). Problem-oriented training promotes spontaneous analogical transfer: Memory-oriented training promotes memory for training. *Memory and Cognition, 19*(6), 543–557.

Osman, M. E., & Hannafin, M. J. (1994). Effects of advance questioning and prior knowledge on science learning. *The Journal of Educational Research, 88*(1), 5–13.

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: Cognitive Science Society.

Roll, I., Aleven, V., & Koedinger, K. R. (2011a). Outcomes and mechanisms of transfer. 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.

Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2007). Designing for metacognition—applying cognitive tutor principles to the tutoring of help seeking. *Metacognition and Learning, 2*(2), 125–140.

Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011b). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction, 21*, 267–280.

Schoenfeld, A. H. (1987). What's all the fuss about metacognition? *Cognitive Science and Mathematics Education* (pp. 189–215). Hillsdale: Erlbaum.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 334–370). New York: MacMillan.

Schwartz, D. L., & Bransford, J. D. (1998). A time for telling. *Cognition and Instruction, 16*(4), 475–522.

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.

Schwartz, D. L., Martin, T., & Pfaffman, J. (2005). How mathematics propels the development of physical knowledge. *Journal of Cognition and Development, 6*(1), 65–88.

Schwartz, D. L., Sears, D., & Chang, J. (2007). Reconsidering prior knowledge. In M. C. Lovett & P. Shah (Eds.), *Thinking With Data* (pp. 319–344). New York: Routledge.

Siegler, R. S. (2002). Microgenetic studies of self-explanation. In N. Granott & J. Parziale (Eds.), *Microdevelopment—Transition Processes In Development And Learning* (pp. 30–58). Cambridge: Cambridge University Press.

Star, J. R., & Rittle-Johnson, B. (2009). It pays to compare: An experimental study on computational estimation. *Journal of Experimental Child Psychology, 102*(4), 408–426.

Taylor, J. L., Smith, K. M., van Stolk, A. P., & Spiegelman, G. B. (2010). Using invention to change how students tackle problems. *CBE Life Sciences Education, 9*(4), 504–512.

Terwel, J., van Oers, B., van Dijk, I., & van den Eeden, P. (2009). Are representations to be provided or generated in primary mathematics education? Effects on transfer. *Educational Research and Evaluation, 15*(1), 25–44.

Tobias, S., & Duffy, T. M. (2009). *Constructivist Instruction: Success or Failure?* New York: Taylor and Francis.

Westermann, K., & Rummel, N. (2012). Delaying instruction—evidence from a study in a university relearning setting. *Instructional Science*. doi:10.1007/s11251-012-9207-8.

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction, 16*(1), 3–118.

Wiedmann, M., Leach, R. C., Rummel, N., & Wiley, J. (2012). Does group composition affect learning by invention? *Instructional Science*. doi:10.1007/s11251-012-9204-y.