CrossMark

ARTICLE

Help Helps, But Only So Much: Research on Help Seeking with Intelligent Tutoring Systems

Vincent Aleven [1] · Ido Roll [2] · Bruce M. McLaren [1] ·
Kenneth R. Koedinger [1]

Abstract Help seeking is an important process in self-regulated learning (SRL). It may influence learning with intelligent tutoring systems (ITSs), because many ITSs provide help, often at the student's request. The Help Tutor was a tutor agent that gave in-context, real-time feedback on students' help-seeking behavior, as they were learning with an ITS. Key goals were to help students become better self-regulated learners and help them achieve better domain-level learning outcomes. In a classroom study, feedback on help seeking helped students to use on-demand help more deliberately, even after the feedback was no longer given, but not to achieve better learning outcomes. The work made a number of contributions, including the creation of a knowledge-engineered, rule-based, executable model of help seeking that can drive tutoring. We review these contributions from a contemporary perspective, with a theoretical analysis, a review of recent empirical literature on help seeking with ITSs, and methodological suggestions. Although we do not view on-demand, principle-based help during tutored problem solving as being as important as we once did, we still view it as helpful under certain circumstances, and recommend that it be included in ITSs. We view the goal of helping students become better self-regulated learners as one of the grand challenges in ITSs research today.

✉ Vincent Aleven
aleven@cs.cmu.edu

Ido Roll
ido.roll@ubc.ca

Bruce M. McLaren
bmclaren@cs.cmu.edu

Kenneth R. Koedinger
koedinger@cmu.edu

[1] Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, USA

[2] Centre for Teaching, Learning, and Technology, University of British Columbia, Vancouver, Canada

🖄 Springer

Introduction

In this commentary we reflect on our article "Toward Meta-cognitive Tutoring: A Model of Help Seeking with a Cognitive Tutor" (Aleven et al. 2006b) and our subsequent work. This line of research is part of an important area within the field of Artificial Intelligent in Education (AIED), which focuses on how learners regulate their own learning and how ITSs can support them (a) in doing so and (b) in becoming better at it. A second goal in the current commentary is to look at recent developments in this area and highlight important questions to address and methodologies as we move forward.

Our 2006 IJAIED paper was published at the midpoint of our project that investigated (a) whether and how an ITS might support effective help seeking and (b) whether student learning would improve as a result. Under a Vygotskian perspective, help seeking can be a step towards being able to operate independently in a given task domain. Help seeking may play an important role in learning with ITSs. As discussed in VanLehn's (2006) excellent cataloging of ITS features, help is a common feature of the "inner loop" of ITSs. Many ITSs provide help in the context of step-based problem solving, often at the student's request, including Andes (VanLehn et al. 2005), Assistments (Heffernan and Heffernan 2014; Razzaq and Heffernan 2010), Cognitive Tutors (Anderson et al. 1995; Koedinger and Corbett 2006), example-tracing tutors (Aleven et al. 2009), Sherlock (Katz et al. 1998; Lesgold et al. 1992), and Wayang Outpost (Arroyo et al. 2014). Other ITS projects have taken a different approach to supporting help seeking, for example, supporting peer tutoring (Walker et al. 2014) or brokering peer help by means of student modeling and reputation systems (Vassileva et al. 2016). In the current article, we focus on *on-demand, principle-based help*. Help of this type is given, by the system, at the student's request (e.g., Razzaq and Heffernan 2010). It states what to do next and why that is a good thing to do, explained in terms of underlying problem-solving principles. Hints of this type are assumed to help students enhance their understanding of key concepts and principles (e.g., Aleven 2013; Anderson 1993) and reduce floundering during problem solving (Aleven and Koedinger 2000). Help seeking is important in many learning contexts other than ITSs, including social contexts, although our project focused on help use within ITSs. It is an important open question how ITSs can help students acquire skills in seeking and using help effectively that transfer to other environments.

We start out by briefly reviewing our prior work on creating and evaluating the Help Tutor, a tutor agent that provided real-time, in-context feedback on students' help-seeking behavior, as they work with an ITS. A key goal was to help students learn more when working with an ITS. We then synthesize relevant literature, both empirical and theoretical, that has accumulated since our 2006 paper. We discuss a number of key issues. First: To what degree does help help in ITSs? That is, what evidence is there that the help offered by an ITS has a beneficial influence on student learning? This question – and the title of the current article – is a nod to a paper by Beck et al. (2008), who conclude, in line with our earlier viewpoint, that

the answer is "yes." Our theoretical analysis and review of the empirical literature indicates that this viewpoint is still tenable today, but also that additional nuance is needed. We also discuss to what extent new research forces us to revise the model of help seeking presented in the IJAIED 2006 paper, a key theoretical contribution of our work. We look at design aspects of providing help and supporting sense making during learning with an ITS. Finally, we present a framework for thinking about the multiple evaluation goals that might be addressed in research on supporting SRL in ITS.

Recap of Main Results and Contributions

The work started with a discovery in student log data from the Geometry Cognitive Tutor in an early educational data mining study (EDM) (Aleven and Koedinger 2000; 2001). Students frequently used the tutor's on-demand help facilities in ways that seemed unlikely to help them learn. For example, they often clicked through the hint levels really quickly to get to the last level (the "bottom-out hint"), which gave the answer. That is, they viewed 68 % of the hint levels prior to the last for less than 1 second (Aleven and Koedinger 2001), not enough time to read and understand the content. Also, students often appeared reluctant to ask for hints. For example, even after 3 errors on a step, the student's next action was a hint request only 34 % of the time (2000). Further, we found a negative correlation between the frequency with which students used the tutor's on-demand help and their learning gains, the latter measured outside the tutor by pre-test/post-test scores. Although this negative correlation no doubt reflects a selection effect – that hints are a sign of struggle during the learning process, which is often associated with lower learning gains (Corbett and Anderson 1995; Wood and Wood 1999) – the finding also suggests that any positive effect on learning from hints is not strong, at least not strong enough to offset the selection effect. The finding of widespread ineffective help seeking was at odds with (in retrospect, idealized) notions at the time of how learners approach learning with an ITS, for example, that they continuously and carefully monitor their comprehension, seeking help when things are not clear (Aleven et al. 2003; Anderson 1993).

These findings led to the idea of extending the tutor so that it helps students learn to use on-demand help more effectively. We decided to focus on *tutoring* as a way of addressing maladaptive help seeking because we had come to believe that many students, as they were working with the Cognitive Tutor software, genuinely did not know when to ask for help or how to use help effectively. We saw an opportunity to improve the effectiveness of tutoring technology by pushing it in a new direction – augmenting the focus on domain-level learning with a new focus on helping students become better learners, focused on self-regulation.

Thus, we built a tutor agent that provides adaptive tutoring on help seeking. We used the same approach and technology that had been used to build Cognitive Tutors for domains like computer programming and mathematics (Anderson et al. 1995), namely, cognitive modeling and model tracing (e.g., Aleven 2010). We first created a rule-based model that captures a range of adaptive and non-adaptive help-seeking behaviors in the context of learning with an ITS. We took a knowledge engineering approach to modeling, through theoretical and empirical cognitive task analysis (e.g., Lovett

1998), guided by analysis of tutor log data and pre/post test data (Aleven et al. 2006b). As emphasized by current theories of SRL, the model was "contextual" in that it captured conditions under which help use is or is not appropriate. It also modeled, in a simple way, students' judgment whether the help received was sufficiently helpful to proceed with problem solving. In its final incarnation, the model comprised approximately 80 production rules. It included a taxonomy of maladaptive help-seeking behaviors with broad categories such as Help Abuse, Help Avoidance, Try-step Abuse, and more detailed subcategories.

We then created a tutor agent, called the Help Tutor, that provides real-time feedback on help seeking, as students work with a step-based ITS. This tutor agent employed the help-seeking model using the standard model-tracing algorithm (e.g., Anderson et al. 1995). We integrated it with the Geometry Cognitive Tutor, so that students would get tutored both on geometry and on help seeking (Aleven et al. 2005). For example, if a student tried clicking through hint levels quickly, the Help Tutor would say: "Try to take more time to read the hint." If the student made multiple errors on a step without asking for a hint, the Help Tutor would say: "Repeated errors may mean that you are not learning, and that you might need some help. Perhaps ask for a hint?" We conducted two classroom studies to test how this contextual feedback on help seeking would influence students' help seeking and learning with the Geometry Cognitive Tutor. We found that this feedback led to a *lasting* improvement in help-seeking behavior, even months after the Help Tutor was turned off. Specifically, students who had received feedback on help seeking used the hints more deliberately: They spent more time per hint level and when they asked for help, requested fewer hint levels. However, we did not find improved domain-level learning due to feedback on help seeking. In a recent data mining study (Roll et al. 2014a), discussed in more detail below, we found evidence in the tutor log data suggesting that, when students have a medium level of skills, hints do have beneficial effect on student learning within the tutor. Specifically, help use led to better performance on the next opportunity to apply the same skill.

The project made a number of contributions to the ITS literature. First, we view the production rule model of help seeking as a research contribution in itself. It is different from prior models of help seeking (e.g., Newman 1994) in that it is detailed and executable, though narrower in scope, as it focuses on help seeking in the context of tutored problem solving. Also, it includes a taxonomy of maladaptive help-seeking behaviors. Operationalizing theoretical constructs is a challenging yet essential endeavor. This work built on prior proposals to use production rules as a way to model SRL processes (Winne and Hadwin 1998). To the best of our knowledge, however, prior to ours, no running model of SRL processes had been created or demonstrated to be useful. As a second key contribution, the project demonstrated the feasibility of using standard ITS technology to provide feedback on a particular SRL skill. While the Help Tutor was not the first ITS that supported some aspect of SRL – e.g., projects by Conati and VanLehn (2000) and Aleven and Koedinger (2002) on supporting self-explanation came earlier – it may have been the first occasion that an ITS provided feedback on how students were self-regulating their learning. Third, it may have been the first ITS project to test whether support for SRL results in a *lasting* improvement in self-regulatory skill.

The main disappointment was that despite improvement in help-seeking behavior, the Help Tutor had no influence on students' domain-level learning outcomes. It is perhaps worth noting that at least one related project found, in a close parallel to our findings, that an intervention aimed at improving help seeking had an influence on the targeted learning behaviors but not on domain-level learning outcomes (Tai et al. 2013). After years of agonizing and soul searching, we have come to view this null result as interesting and important in its own right, contributing to our field's understanding of the role of on-demand help in ITS. We have thought long and hard about what this result might mean. Perhaps, in spite of our best efforts, the model of help seeking was not fully accurate or complete. Possibly, the tutor's help messages should have been more clear in one way or another, or more polite or direct. Perhaps real-time feedback tutoring on help seeking leads to cognitive overload; perhaps post-hoc review of help-seeking behavior avoids such load effectively. Perhaps a wider range of SRL skills (not just help seeking) needs to be supported before we see a measurable impact on student learning. Perhaps students motivations to use or not use help should be addressed. Our experimental design does allow us to tease apart these possible factors. It is likely that many or all these factors contributed to not finding that improved help seeking led to greater pre/post learning gains. In the next sections, we take a step back to look at what other literature has found since.

Does Help Help? Theory on Verbal Sense Making During Problem Solving

As mentioned, on-demand, principle-based hints are common in ITSs that support step-based problem solving (many ITSs do, see VanLehn 2006), so questions regarding their effectiveness are of general import to the field. Hints of this type explain – usually over multiple "levels" of hints – which problem-solving principle applies in the next step of a problem, what the principle says, how it applies, and what concretely to do. Generally, hint sequences of this type tend to be challenging pieces of domain-specific text, especially for novices in a domain, no matter how hard we (as tutor developers) try to make them easily digestible. For students to make sense of these hints requires effort and sophisticated domain-specific (e.g., mathematical) reading comprehension skill. The educational psychology and cognitive science literatures support the notion that learning from instructional explanations can be effective but is challenging (Wittwer and Renkl 2008). Theoretical frameworks in cognitive science suggest a limited but non-negligible role for tutor hints in the context of tutored problem solving in a complex domain. We frame the issue in terms of the Knowledge-Learning-Instruction (KLI) theoretical framework (Koedinger et al. 2012), a recent framework linking cognitive science and educational research to educational practice. Our discussion extends that in Aleven (2013). A fundamental hypothesis in KLI is that the complexity of the instruction needs to match the complexity of the targeted knowledge, in the sense that complex knowledge requires complex instructional processes. Thus, we ask whether tutor hints (as instructional process) match the complexity of the targeted knowledge. We consider both conceptual knowledge of geometry (verbal concepts, theorems, and definitions) and procedural knowledge (i.e., non-verbal problem-solving skill) as potential targets of the instruction. Conceptual knowledge in geometry qualifies as a prime example of complex

knowledge under KLI – verbal knowledge components with a rationale. It therefore requires a complex instructional process that can invoke sense making as a learning mechanism. Under KLI, self-explanation is viewed as such an instructional process. Given that students often do not self-explain spontaneously (Renkl et al. 1998) and given that principle-based hints are not typically accompanied by support for self-explanation, they may not be the best possible way of supporting acquisition of conceptual knowledge. Nonetheless, they may be helpful for some students – those who engage in sense making out of their own volition. It should be noted that theoretical accounts of learning with ITSs do not view hints as the only or even the primary way of learning conceptual knowledge in the context of problem-solving practice. For example, Anderson et al. (1995) recommend (as one of eight Cognitive Tutor principles) presenting "declarative instruction" in the context of problem solving. Although the principle is vague, it suggests that additional instruction (outside of the ITS) on conceptual knowledge is likely to be needed.

Under KLI, procedural knowledge is in a middle stratum of knowledge complexity. (We view procedural knowledge as, in KLI terms, knowledge components with a variable condition and conclusion.) It may therefore be best matched with an instructional process that invokes induction/refinement processes as a learning mechanism. Tutored problem solving (e.g., as supported by an ITS) is such an instructional process, but self-explaining principle-based hints is not. Therefore, tutor hints are not *directly* useful for acquiring procedural knowledge. Under KLI, (verbal) sense making cannot influence (non-verbal) learning *directly*. It can do so *indirectly*, however, for example by providing information that may guide induction/refinement processes that yield procedural knowledge. For example, helping students identify relevant features in problems may help them induce non-verbal rules (i.e., problem-solving knowledge) at an appropriate level of generality, rather than overly specific or overly general versions of such rules. Self-explanation of principle-based hints could have this effect, consistent with the positive effect of verbal self-explanation on the robustness of geometry problem-solving knowledge found by Aleven and Koedinger (2002).

While so far this analysis has focused on explanatory hints, from a theoretical perspective, bottom-out hints (i.e., hints that provide answers, typically found at the end of a hint sequence) might also aid learning of conceptual and procedural knowledge. These hints essentially turn a problem step into a worked example, which could then serve as a basis for analogical problem solving on subsequent opportunities to apply the same knowledge. However, this effect hinges on whether or not students engage in effective sense making (e.g., self-explanation) around bottom-out hints (cf. Shih et al. 2008, discussed below).

In short, our theoretical analysis indicates that principle-based hints during tutored problem solving may aid conceptual learning and may be useful for procedural learning if they help in identifying key problem features. For these beneficial effects to occur, students must self-explain the hints or otherwise make sense of them, which however cannot be taken for granted – students differ in this regard. In short, the theoretical analysis suggests that help helps, but only so much. This analysis represents a shift in our thinking, compared to our viewpoints espoused in the IJAIED 2006 paper, in which we did not focus on the sense making processes that need to occur (and need to be supported) for hints to support learning effectively. We note that this analysis takes a strong *cognitive* perspective on the issue without addressing social, motivational, and

metacognitive aspects of help seeking, which may also have an influence. For example, students' self-assessment of their knowledge or the system's feedback might focus their energy and attention; lack of motivation or a belief that hints are not worth the effort might do the opposite. As another example, help seeking with a tutor might be influenced by the presence of peers in the same computer lab. While the possibility of effects like these does not invalidate our theoretical argument above, it could skew the balance of the different learning mechanisms, resulting in a different analysis.

Does Help Help? Empirical Evidence

Let us now turn to the empirical research regarding how much help helps in learning with an ITS. We consider both experimental studies and data mining studies. Starting with the former, a key way to test the causal hypothesis that principle-based on-demand hints provided by an ITS promote learning is to conduct experimental studies that compare versions of the same tutor that differ only in whether such hints are available. We are not aware of any true experimental studies that did exactly that, although a number of studies approached this ideal quite closely. A study by Anderson and colleagues tested the value of explanatory content in hint and feedback messages in the Lisp Tutor (Anderson et al. 1989). This study compared a version of the tutor that included explanatory content in on-demand hints and system-generated feedback messages against a tutor version that gave only bottom-out hints and only correctness feedback (i.e., did not provide explanatory content). Surprisingly, the explanatory content helped students learn more efficiently, but had no effect on post-test scores. Anderson et al.'s explanation is that even without explanatory content, students were able to construct the explanations for themselves once they knew the correct thing to do, consistent with our theoretical analysis above. This study did not separate the contributions of feedback messages and on-demand hints; also, the hints being tested were not principle-based hints – they explained operators in the programming language LISP. Nonetheless, it is reasonable to assume that these results will generalize to principle-based hints, although only in domains where students tend to be capable of finding explanations for correct steps, once they are given the step. (We do not, however, believe that geometry is such a domain, since the theorems can be complex and non-intuitive). Second, a study by Stamper et al. (2011) compared versions of a tutoring system for logic proof with and without next-step hints. The hints (which were automatically generated from tutor log data) suggested which proof step to try next. The study showed that the hints improve student learning compared to the original system without the hints, which only provided feedback on the correctness of steps. This study suggests that on-demand hints in an ITS can help students learn, but the hints did not mention problem-solving principles – they dealt with search control in logic proof, steering students towards an effective proof process but without explanatory content. Further, the study was not a true experimental study, as there was no random assignment, given the two conditions were run in consecutive semesters. Third, a study by Schworm and Renkl (2006) with a non-intelligent system for studying examples addressed, in a 2 × 2 design, effects of instructional explanations (akin to principle-based hints) and prompted self-explanations (though without feedback). They found that on-demand principle-based help was effective, but not as effective as self-

explanation prompts designed to elicit the same principle-based explanations. In fact, the study showed that this form of on-demand help can undermine the effectiveness of self-explanation prompts, when both prompts and on-demand are available at the same time. In addition, a study by McKendree (1990) varied the *content* of hints and feedback messages (goal feedback versus condition violation feedback) and found that the content had an influence on learning with a geometry tutor. Finally, Razzaq and Heffernan (2010) in a study with the Assistments system found that providing hints on demand (i.e., at the student's request) is more effective (i.e., leads to greater learning gains) than having the system provide them automatically in response to errors. This study shows that the way hints are provided influences learning outcomes and provides a measure of validation for the common practice of providing hints on demand, but it was not fully designed to answer our question, (how much) does help help? As in the McKendree study, it does not fully isolate the effect of hints. In sum, even if none of these studies constitute a fully rigorous confirmation of the beneficial effect of on-demand, principle-based hints in ITSs, these studies strongly suggest that – under certain circumstances – help use may cause learning.

In addition to these experimental studies, some work in EDM has studied help seeking and its relations with learning through analysis of log data from tutoring systems. Typically, these studies look at correlations between measures of help-seeking behavior and learning gains as measured by a pre/post test measuring domain-level learning outside the system. First, a number of studies found positive relations between the frequency of help seeking and learning (Beck et al. 2008; Wood and Wood 1999). (The Wood and Wood paper was a strong influence on our work.) Some work found positive relations not between the "raw" frequency of help seeking and learning, but between the frequency of *adaptive* help seeking and learning. For example, our own IJAIED 2006 paper (Aleven et al. 2006b) found a positive relation between adaptive forms of help seeking and learning, even if the raw frequency of help use (without considering whether it was adaptive or maladaptive) correlated negatively. Similarly, Long and Aleven (2013a) found that although the raw frequency of help use correlated negatively with learning, the time per hint level correlated positively. Arguably, the latter is a measure of deliberate, adaptive help use. These findings underline the importance of keeping *desired* help-seeking behaviors clearly in mind when studying help seeking and its relations with learning.

Some EDM work provides insight regarding *what forms* of help seeking and help use can be effective, and *when* help seeking can be effective. Shih et al. (2008) showed that the time spent with bottom-out hints in an intelligent tutor correlates positively with learning. They interpreted this result as showing that some students take advantage of bottom-out hints as opportunities for spontaneous self-explanation (i.e., unprompted and unsupported by the tutor) similar to self-explaining worked examples, which has been shown to be effective in learning (Chi et al. 1989; Renkl 2013). This reasoning is in line with Anderson et al.'s (1989) observation that students were able to construct effective self-explanations when given the answer. It is also in line with the theoretical analysis of learning from hints presented above. As a counterpoint, Mathews and colleagues (2008) found – analyzing data from constraint-based tutors – that a high rate of bottom-out hints corresponded with less learning, while a high rate of low-to-intermediate hints corresponded with greater learning. As another counterpoint, a number of studies found advantages for worked examples in the context of tutored

problem solving (e.g., McLaren et al. 2016; Razzaq and Heffernan 2009; Salden et al. 2010). This finding strongly suggests that bottom-out hints are not as effective as worked examples. Taken together, the two studies illustrate that much depends on how bottom-out hints are used. Roll et al. (2014a) studied *when* in the learning process hints might be effective. Analyzing log data from the Geometry Cognitive Tutor, they looked at "local" within-tutor learning, specifically, at whether hint use on a step had an effect on the next opportunity to apply the same knowledge component. They found that hints are beneficial for local within-tutor learning when students have a medium level of skill. When students have a low or high level of skill, attempts at solving (without help) are more effective. These findings suggest, in terms of KLI (see our analysis above) that non-verbal knowledge built up through problem-solving practice might help prepare students for principle-based hints, and that principle-based hints may help primarily in refining knowledge obtained from initial problem-solving attempts. They are a key piece of evidence in our own study that help does help, though only in certain circumstances (when the skill level is medium) and only to a limited degree (the effect was seen in tutor data but apparently was not strong enough to be measurable at post test).

Finally, recent EDM work has shown that there are individual differences in how students utilize help and how it helps them during tutor problems (Goldin et al. 2012, 2013). Although the work did not show that individual differences in hint processing are associated with differential learning gains, it does highlight the self-regulatory character of hint use in an ITS and adds to an interesting line of inquiry within EDM regarding individual differences in students' self-regulation of learning.

So, does help help? Compared to our thinking in 2006, we are now much more cognizant of the fact that learning from on-demand, principle-based next-step hints is very challenging for novice students. Cognitive theory such as KLI reserves a limited (but non-negligible) role for sense-making processes in the context of tutored problem solving. Experimental research and correlational data mining studies provide evidence that is in line with this theoretically-derived view. They provide some evidence that hints can be helpful and *when* hints may be useful, namely, at an intermediate stage in the learning process. At least one data mining study suggests the importance of careful sense making (e.g., self-explanation). In sum, these results and analyses suggest that under appropriate circumstances, (verbal) principle-based hints provide important opportunities for sense making and learning, even if they are challenging opportunities and even if the effect is perhaps small, compared to the induction and refinement of problem-solving knowledge that happens through problem solving practice with feed-back. In other words, help helps, but only so much …

Understanding and Modeling Help Seeking with an ITS

As mentioned, we view the model of help seeking presented in our 2006 IJAIED paper (Aleven et al. 2006b) as a theoretical contribution, because it makes explicit, for inspection, discussion, and empirical verification, how students should seek help with an ITS in ways that positively affect their learning. In this section, we consider the questions: What does the model of help seeking teach us that is still relevant today? What additional insight does later research add? We also look at how students' actual help-seeking behavior tends to conform or differ from that prescribed by the model.

An empirical finding that is still largely intact today is that Help Abuse, a particular form of maladaptive help seeking, tends to be frequent. By Help Abuse we mean help use that avoids careful reading and sense making, for example, clicking through hints while spending very little time reading them, and then copying the answer provided in the bottom-out hint. The analysis presented in our IJAIED 2006 shows this behavior to be both frequent and negatively correlated with learning, as measured by pre/post tests outside the tutor. We saw frequent Help Abuse in other tutor units as well (e.g., Aleven et al. 2006a), although it was not quite as frequent as that reported in the IJAIED paper. Work on "gaming the system" (Baker et al. 2008a, b, 2013) also found that Help Abuse (one of two key categories of gaming the system) is frequent. On the other hand, other work suggests that Help Abuse (as we defined it originally) may not always be detrimental for learning. A study by Shih et al. (2008) found a positive relation between time spent with bottom-out hints and learning, which they interpreted as suggesting that students self-explain bottom-out hints similar to steps in a worked example. Extending that argument somewhat beyond Shih et al.'s (2008) analysis, perhaps skipping inter-mediate hints is not detrimental if students use the ensuing bottom-out hint as oppor-tunity for sense making. We note that this explanation again emphasizes the importance of sense making, as discussed above.

Our understanding of Help Avoidance, another form of maladaptive help seeking, has been refined by later work, including our own. Initially, we defined Help Avoidance as not seeking help when help use would appear to be beneficial for learning, for example, when a step was not familiar (detected through self-assessment) or when it is not clear how to fix an error (detected through self-assessment, or as evidenced by errors on a step). Our initial findings were that Help Avoidance is frequent and, at least in one of our analyses, associated with poorer learning. However, later data mining work by Shih et al. (2010) suggested that persisting in attempts at solving tends to be more effective than requesting help. Our own recent work Roll et al. (2014a) indicates that hints help learning when students have a medium level of mastery of the given knowledge components. This finding suggests – although without definitively proving it – that Help Avoidance may be an issue only in this intermediate stage of learning. It also refines an assumption in our model of help seeking, that students closer to mastering a given skill component are thought to be better off with less help (i.e., fewer hint levels) regarding that skill component. To counterbalance these findings, however, a recent paper by Baker et al. (2011) looked at using tutor log data to predict student performance on a transfer posttest. Help Avoidance had a negative relationship with performance on the transfer test. We do not quite know how to reconcile these findings, so let us just say that our community's understanding of how Help Avoidance affects learning is still evolving.

Finally, somewhat informally, students' actual help-seeking behavior appears not to align with the model's basic assumption that students should self-assess their mastery of the knowledge needed for the problem step at hand as a basis for deciding whether to seek help. Under the model, students should continuously ask themselves: Is this step familiar? If not, let me ask for a hint right off the bat. Do I understand how I can fix my error? If so, I should just fix it; otherwise I should request a hint. Now that I have read this hint level, do I know enough to try the step, or do I need a more detailed hint? In an ITS, self-assessment of this kind may be aided by the tutor's open learner model (Bull and Kay 2007, 2010; Dimitrova 2003; Long and Aleven 2013b; Mitrovic and Martin

2007). Although we do not have direct data on how frequently students were self-assessing their knowledge, it is our strong impression that the decision to seek help is made more often on the basis of error feedback from the tutor. For example, students rarely ask for a hint before attempting a step. Interestingly, in one of our studies, we have seen that interventions aimed at supporting self-assessment had the result that students used help more deliberately (e.g., Long and Aleven 2013a).

How to Design Help that Helps?

We now turn to design implications. First, in light of the findings presented above, should ITSs provide on-demand principle-based hints? Or is it sufficient that they help students learn targeted problem-solving principles inductively (Koedinger et al. 2012), through tutored problem solving with feedback but without verbal sense making (cf. Aleven 2013). Our recommendation, based on the research discussed above, is that ITS developers continue to provide principle-based hints. These hints may support useful sense making, even if their effectiveness may vary by domain and by student, and even if, for any given student, hints may be effective only some of the time. In addition, it seems clear also that ITSs need bottom-out hints – if nothing else, they help students get unstuck when they are stuck. In addition, bottom-out hints present useful opportunities for unprompted, spontaneous self-explanation (Shih et al. 2008). Self-explanation is associated with enhanced learning outcomes (Chi et al. 1989), even if not all students benefit equally (Renkl et al. 1998).

If hints are needed, what are good ways of structuring/writing them? Hints should abstractly (but succinctly) characterize the problem-solving knowledge (Anderson et al. 1995), by stating in general terms both the action to be taken, the conditions under which this particular action is appropriate, and the domain principle (e.g., geometry theorem) that justifies the action. Hints also need to explain briefly how the problem-solving principle applies to the given problem step. It may be clear that these requirements are not easy to meet with simple, concise language. Often, this advice conflicts with keeping hints short. That is, writing good hints is a balancing act.

Given that learning from principle-based hints is hard, however, it is important that ITS research looks for better ways of supporting sense making in and around tutors. A number of researchers have suggested that self-explanation prompts should replace tutor hints (except the bottom-out hint). This interesting idea finds some support in the data mining study by Shih et al. (2008), discussed above. However, we are not aware of any experimental studies that rigorously tested this idea. A study by Butcher and Aleven (2013) focused on self-explaining steps in response to errors, rather than through hints. They demonstrated, in a study with the Geometry Cognitive Tutor, that prompting for self-explanations after errors in tutor problems enhances student learning. Specifically, students were prompted to select the relevant diagram elements in an interactive diagram, as a form of visual self-explanations. The same idea could be applied to hints as well. That is, prompted self-explanations with feedback could supplement hints and perhaps even replace them entirely. Some studies suggested benefits of making hints interactive, although they also point to interactions of interactivity with learner characteristics. Razzaq and Heffernan (2006) found evidence suggesting that breaking down a problem into steps is more effective than providing a long hint. Arroyo et al. (2000) found that

elementary school girls learn better with highly interactive hints whereas elementary school boys do better with hints that are not highly interactive.

We also recommend more fully exploring support for self-explaining bottom-out hints, for example, with richer self-explanations (Aleven et al. 2004; Butcher and Aleven 2013), or to bring together lines of research on ITS for reading comprehension (Jackson and McNamara 2013) and help seeking with ITSs. Further, it is probably very useful if some of the sense making happens outside of the tutor, conform the Cognitive Tutors principle to "provide instruction in the context of problem solving" (Anderson et al. 1995). Finally, it may be interesting to investigate whether it is effective for ITSs to provide different hints to students depending on their level of knowledge of the targeted problem-solving principle. A simple form of doing so is through "contingent tutoring," a notion pioneered by Wood and Wood (1999), in which more abstract hint levels are gradually introduced as students gain competence.

Methodological Aspects

We discuss two distinctive methodological aspects of the work. First, the work is somewhat unusual within the field of AIED in that it used a knowledge engineering approach to model aspects of SRL. By "knowledge engineering" we mean that we engineered the model top down, guided by theories of help seeking and aided substantially by data, including informal classroom observations and the mining of log data and pre/post-test data. Others have applied machine learning techniques, such as methods for building classifiers, Markov Processes, sequence mining techniques, cluster analysis, and so forth (Bouchet et al. 2013; Conati and Kardan 2013; Kinnebrew et al. 2014; Montalvo et al. 2010; Sabourin et al. 2013). A key advantage of a knowledge engineering approach is that it results in an *interpretable* and *explainable* model; that is, a model described in terms humans can understand. Such a model can convey novel insights into the modeled phenomena and can therefore be viewed as a form of theory formation (e.g., Aleven 2013). On the other hand, highly parameterized machine-learned models tend not to be interpretable. Although they provide potentially useful black boxes to be plugged into systems, they do not advance our scientific understanding nor provide an avenue for application of insights other than by using the model as a black box (cf., Liu et al. 2014). We foresee that in the future, more encompassing models of SRL will be needed in advanced learning technologies, given that SRL theoretical frameworks integrate a very wide range of cognitive, metacognitive, motivational, and social processes (Pintrich 2004; Winne and Hadwin 1998; Zimmerman 2000, 2008). Support for these processes could benefit from both types of modeling approaches.

A second methodological aspect of the work on help seeking is that it evaluated whether there were lasting effects of support for SRL, similar to Bransford and Schwartz' (1999) framework of preparation for future learning (see also White and Frederiksen 1998). Generally, in thinking about effects that support for SRL might have on students, it may help to make four key distinctions (cf. Koedinger et al. 2009; Roll et al. 2014b).

1. Effects an intervention may have on targeted SRL processes (e.g., help seeking, self-assessment, and so forth) versus on domain-level learning.

2. Effects while the intervention is in effect ("current learning") versus learning that comes afterwards ("future learning").
3. Effects on learning in the same learning environment in which the SRL intervention was embedded versus in a new learning environment.
4. Effects on learning in the same domain versus a different domain.

These four items define a space of 16 possible effects on learning that interventions to support SRL may have. This framework may help researchers evaluate empirical research on SRL (in ITS but also in other learning environments), and may help define goals for this type of research at various levels of ambition.

We relate the evaluation studies of the Help Tutor to this framework. In the classroom studies with the Help Tutor, we addressed two key evaluation questions: How does feedback on help seeking influence students' geometry learning while the feedback is in effect (i.e., how does it influence their *current domain-level learning*)? How does it influence students' help-seeking behavior both during and after the feedback (i.e., how does it their *current* and *future SRL – specifically, help seeking*)? Regarding the first question, we assessed students' domain-level learning while the intervention was in effect using a pre/post-test of geometry skill. Regarding the second question, we assessed whether tutoring on help seeking had an *immediate* effect on students' help-seeking behavior by analyzing tutor log data collected during the experiment (i.e., while the Help Tutor was in effect). In addition, we measured whether tutoring help seeking had a *lasting* effect on students' help-seeking behavior in three ways: using hypothetical help-seeking scenarios administered as part of the post-test, with post-test geometry items that featured embedded on-demand help, and by analyzing students' help-seeking behavior in tutor units that followed the ones in which the Help Tutor was used. This approach, in which aspects of students' SRL are measured both during and after the SRL intervention is over, was quite novel at the time within the field of AIED. It is still rare today (but see Leelawong and Biswas 2008), but in our opinion it is key to moving the area of SRL and ITS forward.

As perhaps a glimpse of what future research on SRL with ITSs might bring, we consider, again within the framework presented above, what additional aspects of student learning could have been evaluated in our classroom studies but were not. First, it would have been interesting to measure whether there was any effect on students' *future domain-level learning within the same environment*, for example in later tutor units in the Geometry Tutor. Doing so was not practical in our project for a variety of reasons. Nonetheless, we see it as a key challenge for the field of AIED research to understand under what circumstances interventions focused on supporting SRL lead to improved future domain-level learning. Second, it would be interesting to study whether students' *future help seeking and learning* would be affected *when working on new content in the same environment* (e.g., Cognitive Tutor Algebra II, say). As a third example, it is an important question whether improved help seeking within an ITS (such as that shown within the Help Tutor project) results in *improved help seeking within other environments*, such as classrooms or small-group project-based learning. This transfer cannot be taken for granted (and quite possibly, more is needed than the Help Tutor to achieve such transfer). This is only a sampling from the 16 evaluation possibilities defined by the framework. They point to ambitious, interesting research questions. More generally, we hope the framework will provide useful guidance for ITS researchers and SRL researchers.

Discussion and Conclusion

The main contributions of the work on the Help Tutor were an executable, rule-based model of help seeking during tutored problem solving, a demonstration that feedback on help seeking is feasible with standard tutoring technology, and an experimental classroom study showing that such feedback can lead to improved "local domain-level learning" and a lasting increase in how deliberately students use help, although the hypothesis that this feedback would lead to improved out-of-tutor transfer of domain-level learning was not confirmed. These results are relevant to all tutors that provide on-demand help to students, if which there are many. Key methodological features of the work were a data-driven knowledge engineering approach to modeling SRL skills (to be contrasted with machine learning approaches) and experimentally testing effects on students' *future* SRL. The current article reviews the role of help seeking with ITS through a theoretical analysis and a review of empirical research. We discuss implications for the design of ITS and methodological issues, and we point to important research directions.

Our understanding of the role of help seeking as part of learning through tutored problem solving has evolved substantially since our 2006 IJAIED paper. Our reading of the empirical literature is that hints and help seeking can help students learn, but that they are not the strong influence on student learning that we once thought they were. Hints help some of the time, namely, when students have an intermediate level of skill (Roll et al. 2014a). (A study that tried to replicate this result in other data sets would be highly valuable.) Also, hints and help seeking help only so much, based both on our theoretical analysis based on the KLI framework and our reading of the empirical studies. We recommend that ITS developers continue to include principle-based hints in tutors. Bottom-out hints are useful, too. They may help students to continue when stuck, and may help learning when spontaneously self-explained (Shih et al. 2008). However, it is clear that spontaneous self-explanation does not occur often. Therefore, it will be helpful to design and evaluate support for explaining bottom-out hints or other forms of sense making around hints.

The model of help seeking presented in our 2006 IJAIED paper continues to provide a useful foundation for understanding help seeking, but some modifications of the model are in order, based on recent research. For example, not captured in the model is that bottom-out hints may aid learning, provided students self-explain them (Shih et al. 2008). Also, the finding that hints are primarily helpful for learning when the student has a medium level of skill (Roll et al. 2014a) means that the model should prescribe help use in a narrower range of circumstances.

Although by now, the area of help seeking and help use with ITSs has been quite thoroughly researched, many interesting open research questions remain. How can the sense-making processes required for learning from hints be supported? Further, motivational aspects of help seeking have not been researched thoroughly enough. It is likely that some examples of poor help seeking, such as Help Abuse, are due to a lack of student motivation. For papers dealing with motivational and social aspects of help seeking in the context of advanced learning technologies, see Baker et al. 2008b; Howley et al. 2014; Tai et al. 2013.

We continue to subscribe to the viewpoint that metacognition and self-regulation are key factors in students' learning and their learning outcomes and that these abilities are

amenable to instruction, at least to a degree. We see continued value in investigating how ITS methodologies and technologies can best be used to support SRL and with what effect on student learning. This viewpoint is shared by many other researchers, as there is much current AIED/ITS research around this question. Our hunch is that it will be useful to take a more comprehensive view of self-regulation, as a number of researchers have started to do, and study multiple SRL constructs in conjunction, consistent with SRL theories (Azevedo et al. 2012; Kinnebrew et al. 2013; Long and Aleven 2013b; Poitras and Lajoie 2013; Roll et al. 2012; Sabourin et al. 2013). In doing so, it is important to take into account not just the cognitive and metacognitive aspects of self-regulated learning, but also the motivational aspects. Existing SRL theoretical frameworks may be useful in this regard, as they often encompass a broad range of motivational constructs.

Methodologically, it is important to study effects of SRL interventions at multiple levels: effects while the intervention is in effect versus effects that occur afterwards; domain-level learning outcomes versus improvements in SRL; effects in the same learning environment versus a different one; and regarding task domain. The questions of whether and how ITS can help learners acquire robust, lasting SRL skills that transfer to new learning environments remain wide open. Tackling these questions looms as an important challenge for our field.

References

Aleven, V. (2010). Rule-based cognitive modeling for intelligent tutoring systems. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (Eds.), *Studies in Computational Intelligence: Vol. 308. Advances in intelligent tutoring systems* (pp. 33–62). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-14363-2_3.

Aleven, V. (2013). Help seeking and intelligent tutoring systems: Theoretical perspectives and a step towards theoretical integration. In: Azevedo, R., & Aleven, V. (Eds.), *International Handbook of Metacognition and Learning Technologies* (pp. 311–335). New York: Springer. doi: 10.1007/978-1-4419-5546-3_21.

Aleven, V., & Koedinger, K. R. (2000). Limitations of student control: Do students know when they need help? In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), *Proceedings of the 5th International Conference on Intelligent Tutoring Systems, ITS 2000* (pp. 292–303). Berlin: Springer.

Aleven, V., & Koedinger, K. R. (2001). Investigations into help seeking and learning with a Cognitive Tutor. In R. Luckin (Ed.), *Papers of the AIED-2001 workshop on help provision and help seeking in interactive learning environments* (pp. 47–58). Retrieved from http://www.sussex.ac.uk/Users/bend/aied2001/helpworkshopnode7.html.

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.

Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research, 73*(3), 277–320.

Aleven, V., Ogan, A., Popescu, O., Torrey, C., & Koedinger, K. (2004). Evaluating the effectiveness of a tutorial dialogue system for self-explanation. In J. C. Lester, R. M. Vicario, & F. Paraguaçu (Eds.), *Proceedings of 7th International Conference on Intelligent Tutoring Systems, ITS 2004* (pp. 443–454). Berlin Heidelberg: Springer. doi:10.1007/978-3-540-30139-4_42.

Aleven, V., Roll, I., McLaren, B. M., Ryu, E. J., & Koedinger, K. R. (2005). An architecture to combine meta-cognitive and cognitive tutoring: Pilot testing the Help Tutor. In C. K. Looi, G. McCalla, B. Bredeweg, & J. Breuker (Eds.), *Proceedings of the 12th International Conference on Artificial Intelligence in Education, AIED 2005* (pp. 17–24). Amsterdam: IOS Press.

Aleven, V., McLaren, B. M., & Koedinger, K. R. (2006a). Towards computer-based tutoring of help-seeking skills. In S. Karabenick & R. Newman (Eds.), *Help Seeking in Academic Settings: Goals, Groups, and Contexts* (pp. 259–296). Mahwah: Erlbaum.

Aleven, V., McLaren, B. M., Roll, I., & Koedinger, K. R. (2006b). Toward meta-cognitive tutoring: A model of help seeking with a cognitive tutor. *International Journal of Artificial Intelligence in Education, 16*, 101–128.

Aleven, V., McLaren, B. M., Sewall, J., & Koedinger, K. R. (2009). A new paradigm for intelligent tutoring systems: Example-Tracing tutors. *International Journal of Artificial Intelligence in Education, 19*(2), 105–154.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale: Lawrence Erlbaum Associates.

Anderson, J. R., Conrad, F. G., & Corbett, A. T. (1989). Skill acquisition and the LISP tutor. *Cognitive Science, 13*(4), 467–505. doi:10.1016/0364-0213(89)90021-9.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive Tutors: Lessons learned. *The Journal of the Learning Sciences, 4*(2), 167–207.

Arroyo, I., Beck, J., Woolf, B. P., Beal, C. R., & Schultz, K. (2000). Macroadapting animalwatch to gender and cognitive differnces with respect to hint interactivity and symbolism. In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), *Proceedings of the 5th International Conference on Intelligent Tutoring Systems, ITS 2000* (pp. 574–583). Berlin: Springer Verlag. doi:10.1007/3-540-45108-0_61.

Arroyo, I., Woolf, B. P., Burleson, W., Muldner, K., Rai, D., & Tai, M. (2014). A multimedia adaptive tutoring system for mathematics that addresses cognition, metacognition and affect. *International Journal of Artificial Intelligence in Education, 24*(4), 387–426. doi:10.1007/s40593-014-0023-y.

Azevedo, R., Landis, R. S., Feyzi-Behnagh, R., Duffy, M., Trevors, G., Harley, J. M., Hossain, G. (2012). The effectiveness of pedagogical agents' prompting and feedback in facilitating co-adapted learning with metatutor. In S. A. Cerri, W. J. Clancey, G. Papadourakis, & K. Panourgia (Eds.), *Proceedings of the 11th International Conference on Intelligent Tutoring Systems, ITS 2012* (pp. 212–221). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-30950-2_27.

Baker, R. S. J. d., Gowda, S. M., & Corbett, A. T. (2011). Towards predicting future transfer of learning. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Lecture Notes in Computer Science: Artificial intelligence in Education: 15th International Conference, AIED 2011* (Vol. 6738, pp. 23–30). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-21869-9_6.

Baker, R. S. J. D., Corbett, A. T., Roll, I., & Koedinger, K. R. (2008a). Developing a generalizable detector of when students game the system. *User Modeling and User-Adapted Interaction, 18*(3), 287–314. doi:10.1007/s11257-007-9045-6.

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008b). Why students engage in "gaming the system" behavior in interactive learning environments. *Journal of Interactive Learning Research, 19*(2), 185–224.

Baker, R. S. J. D., Corbett, A. T., Roll, I., Koedinger, K. R., Aleven, V., Cocea, M., Mathews, M. (2013). Modeling and studying gaming the system with educational data mining. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 97–115). New York: Springer. doi:10.1007/978-1-4419-5546-3_7.

Beck, J. E., Chang, K., Mostow, J., & Corbett, A. T. (2008). Does help help? Introducing the Bayesian evaluation and assessment methodology. In B. Woolf, E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Proceedings of the 9th International Conference on Intelligent Tutoring Systems, ITS 2008* (pp. 383–394). Berlin: Springer.

Bouchet, F., Harley, J. M., Trevors, G. J., & Azevedo, R. (2013). Clustering and profiling students according to their interactions with an intelligent tutoring system fostering self-regulated learning. *Journal of Educational Data Mining, 5*(1), 104–146.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: a simple proposal with multiple implications. *Review of Research in Education, 24*(1), 61–100.

Bull, S., & Kay, J. (2007). Student models that invite the learner in: The SMILI: Open learner modelling framework. *International Journal of Artificial Intelligence in Education, 17*(2), 89–120.

Bull, S., & Kay, J. (2010). Open learner models. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (Eds.), *Studies in Computational Intelligence: Vol. 308. Advances in intelligent tutoring systems* (pp. 301–322). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-14363-2_15.

Butcher, K., & Aleven, V. (2013). Using student interactions to foster rule-diagram mapping during problem solving in an intelligent tutoring system. *Journal of Educational Psychology, 105*(4), 988–1009. doi:10.1037/a0031756.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*(2), 145–182. doi:10.1016/0364-0213(89)90002-5.

Conati, C., & Kardan, S. (2013). Student modeling: Supporting personalized instruction, from problem solving to exploratory open ended activities. *AI Magazine, 34*(3), 13–26.

Conati, C., & Vanlehn, K. (2000). Toward computer-based support of meta-cognitive skills: A computational framework to coach self-explanation. *International Journal of Artificial Intelligence in Education, 11*(4), 389–415.

Corbett, A. T., & Anderson, J. R. (1995). Knowledge tracing: modeling the acquisition of procedural knowledge. *User Modeling and User-Adapted Interaction, 4*(4), 253–278.

Dimitrova, V. (2003). STyLE-OLM: Interactive open learner modelling. *International Journal of Artificial Intelligence in Education, 13*(1), 35–78.

Goldin, I. M., Koedinger, K. R., & Aleven, V. (2012). Learner differences in hint processing. In K. Yacef, O. Zaïane, A. Hershkovitz, M. Yudelson, & J. Stamper (Eds.), *Proceedings of the 5th International Conference on Educational Data Mining (EDM 2012)* (pp. 73–80). Worcester: International Educational Data Mining Society.

Goldin, I. M., Koedinger, K. R., & Aleven, V. (2013). Hints: You can't have just one. In S. K. D'Mello, R. A. Calvo, & A. Olney (Eds.), *Proceedings of the 6th International Conference on Educational Data Mining (EDM 2013)* (pp. 232–235). Worcester: International Educational Data Mining Society.

Heffernan, N. T., & Heffernan, C. L. (2014). The Assistments ecosystem: Building a platform that brings scientists and teachers together for minimally invasive research on human learning and teaching. *International Journal of Artificial Intelligence in Education, 24*(4), 470–497. doi:10.1007/s40593-014-0024-x.

Howley, I., Kanda, T., Hayashi, K., & Rosé, C. (2014). Effects of social presence and social role on help-seeking and learning. In G. Sagerer, M. Imai, T. Belpaeme, & A. Thomaz (Eds.), *HRI '14: Proceedings of the 2014 ACM/IEEE International Conference on Human-robot Interaction* (pp. 415–422). New York: ACM. doi:10.1145/2559636.2559667.

Jackson, G. T., & McNamara, D. S. (2013). Motivation and performance in a game-based intelligent tutoring system. *Journal of Educational Psychology, 105*(4), 1036–1049. doi:10.1037/a0032580.

Katz, S., Lesgold, A., Hughes, E., Peters, D., Eggan, G., Gordin, M., & Greenberg, L. (1998). Sherlock 2: An intelligent tutoring system built on the LRDC framework. In C. P. Bloom & R. B. Loftin (Eds.), *Facilitating the development and use of interactive learning environments* (pp. 227–258). Mahwah: Lawrence Erlbaum Associates.

Kinnebrew, J. S., Loretz, K. M., & Biswas, G. (2013). A contextualized, differential sequence mining method to derive students' learning behavior patterns. *Journal of Educational Data Mining, 5*(1), 190–219.

Kinnebrew, J. S., Segedy, J. R., & Biswas, G. (2014). Analyzing the temporal evolution of students' behaviors in open-ended learning environments. *Metacognition and Learning, 9*(2), 187–215. doi:10.1007/s11409-014-9112-4.

Koedinger, K. R., & Corbett, A. T. (2006). Cognitive Tutors: Technology bringing learning sciences to the classroom. In R. K. Sawyer (Ed.), *The Cambridge Handbook of the Learning Sciences* (pp. 61–78). New York: Cambridge University Press.

Koedinger, K. R., Aleven, V., Roll, I., & Baker, R. (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.), *The Educational Psychology Series: Handbook of Metacognition in Education* (pp. 897–964). London: Routledge Press.

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The Knowledge-Learning-Instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive Science, 36*(5), 757–798. doi:10.1111/j.1551-6709.2012.01245.x.

Leelawong, K., & Biswas, G. (2008). Designing learning by teaching agents: the Betty's Brain system. *International Journal of Artificial Intelligence in Education, 18*(3), 181–208.

Lesgold, A., Lajoie, S., Bunzo, M., & Eggan, G. (1992). SHERLOCK: A coached practice environment for an electronics troubleshooting job. In J. Larkin & R. W. Chabay (Eds.), *Computer-assisted instruction and intelligent tutoring systems: Shared goals and complementary approaches. Technology in education series* (pp. 201–238). Hillsdale: Lawrence Erlbaum Associates, Inc.

Liu, R., Koedinger, K. R., & McLaughlin, E. A. (2014). Interpreting model discovery and testing generalization to a new dataset. In J. Stamper, Z. Pardos, M. Mavrikis, & B. M. McLaren (Eds.), *Proceedings of the 7th International Conference on Educational Data Mining (EDM 2014)* (pp. 107–113). Worcester: International Educational Data Mining Society.

Long, Y., & Aleven, V. (2013a). Skill diaries: Improve student learning in an intelligent tutoring system with periodic self-assessment. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Proceedings of the 16th International Conference on Artificial Intelligence in Education, AIED 2013* (pp. 249–258). Berlin Heidelberg: Springer. doi:10.1007/978-3-642-39112-5_26.

Long, Y., & Aleven, V. (2013b). Supporting students' self-regulated learning with an open learner model in a linear equation tutor. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Proceedings of the 16th International Conference on Artificial Intelligence in Education, AIED 2013* (pp. 249–258). Berlin: Springer.

Lovett, M. C. (1998). Cognitive task analysis in service of intelligent tutoring system design: A case study in statistics. In B. P. Goettl, H. M. Halff, C. L. Redfield, & V. Shute (Eds.) *Intelligent Tutoring Systems, Proceedings of the Fourth International Conference* (pp. 234–243). Lecture Notes in Computer Science, 1452. Berlin: Springer-Verlag.

Mathews, M., Mitrović, T., & Thomson, D. (2008). Analysing high-level help-seeking behaviour in ITSs. In W. Nejdl, J. Kay, P. Pu, & E. Herder (Eds.), *Adaptive Hypermedia and Adaptive Web-based Systems: 5th International Conference, AH 2008* (pp. 312–315). Berlin, Heidelberg: Springer. doi:10.1007/978-3-540-70987-9_42.

McKendree, J. (1990). Effective feedback content for tutoring complex skills. *Human-Computer Interaction, 5*(4), 381–413. doi:10.1207/s15327051hci0504_2.

McLaren, B. M., van Gog, T., Ganoe, C., Karabinos, M., & Yaron, D. (2016). The efficiency of worked examples compared to erroneous examples, tutored problem solving, and problem solving in computer-based learning environments. *Computers in Human Behavior, 55*, 87–99. doi:10.1016/j.chb.2015.08.038.

Mitrovic, A., & Martin, B. (2007). Evaluating the effect of open student models on self-assessment. *International Journal of Artificial Intelligence in Education, 17*(2), 121–144.

Montalvo, O., Baker, R. S. J. D., Sao Pedro, M. A., Nakama, A., & Gobert, J. (2010). Identifying students' inquiry planning using machine learning. In R. S. J. D. Baker, A. Merceron, & P. I. Pavlik Jr. (Eds.), *Proceedings of the 3rd International Conference on Educational Data Mining, EDM 2010* (pp. 141–150). Worcester: International Educational Data Mining Society.

Newman, R. S. (1994). Adaptive help seeking: A strategy of self-regulated learning. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulation of learning and performance: Issues and educational applications* (pp. 283–301). Hillsdale: Lawrence Erlbaum Associates, Inc.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*(4), 385–407.

Poitras, E. G., & Lajoie, S. P. (2013). A domain-specific account of self-regulated learning: the cognitive and metacognitive activities involved in learning through historical inquiry. *Metacognition and Learning, 8*(3), 213–234. doi:10.1007/s11409-013-9104-9.

Razzaq, L., & Heffernan, N. T. (2006). Scaffolding vs. hints in the Assistment system. In M. Ikeda, K. D. Ashley, & T. W. Chan (Eds.), *Proceedings of the 8th International Conference on Intelligent Tutoring Systems, ITS 2006* (pp. 635–644). Berlin, Heidelberg: Springer-Verlag. doi:10.1007/11774303_63.

Razzaq, L. M., & Heffernan, N. T. (2009). To tutor or not to tutor: That is the question. In V. Dimitrova, R. Mizoguchi, B. du Boulay, & A. Graesser (Eds.), *Proceedings of the 14th International Conference on Artificial Intelligence in Education, AIED 2009* (pp. 457–464). Amsterdam: Ios Press.

Razzaq, L., & Heffernan, N. T. (2010). Hints: Is it better to give or wait to be asked? In V. Aleven, J. Kay, & J. Mostow (Eds.), *Lecture Notes in Computer Science: Proceedings of the 10th International Conference on Intelligent Tutoring Systems, ITS 2010* (Vol. 1, pp. 115–124). Berlin: Springer.

Renkl, A. (2013). Toward an instructionally oriented theory of example-based learning. *Cognitive Science, 38*(1), 1–37. doi:10.1111/cogs.12086.

Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: the effects of example variability and elicited self-explanations. *Contemporary Educational Psychology, 23*(1), 90–108.

Roll, I., Holmes, N. G., Day, J., & Bonn, D. (2012). Evaluating metacognitive scaffolding in guided invention activities. *Instructional Science, 40*(4), 1–20. doi:10.1007/s11251-012-9208-7.

Roll, I., Baker, R. S. J. D., Aleven, V., & Koedinger, K. R. (2014a). On the benefits of seeking (and avoiding) help in online problem-solving environments. *Journal of the Learning Sciences, 23*(4), 537–560. doi:10.1080/10508406.2014.883977.

Roll, I., Stampfer Wiese, E., Long, Y., Aleven, V., & Koedinger, K. R. (2014b). Tutoring self- and co-regulation with intelligent tutoring systems to help students acquire better learning skills. In R. Sottilare, A. Graesser, X. Hu, & H. Holden (Eds.), *Design recommendations for adaptive intelligent tutoring systems* (Vol. II, Instructional Management, pp. 169–182). Orlando, FL: US Army Research Laboratory.

Sabourin, J. L., Shores, L. R., Mott, B. W., & Lester, J. C. (2013). Understanding and predicting student self-regulated learning strategies in game-based learning environments. *International Journal of Artificial Intelligence in Education, 23*(1–4), 94–114. doi:10.1007/s40593-013-0004-6.

Salden, R. J. C. M., Koedinger, K. R., Renkl, A., Aleven, V., & McLaren, B. M. (2010). Accounting for beneficial effects of worked examples in tutored problem solving. *Educational Psychology Review, 22*(4), 379–392. doi:10.1007/s10648-010-9143-6.

Schworm, S., & Renkl, A. (2006). Computer-supported example-based learning: when instructional explanations reduce self-explanations. *Computers & Education, 46*(4), 426–445.

Shih, B., Koedinger, K. R., & Scheines, R. (2008). A response time model for bottom-out hints as worked examples. In R. S. J. d. Baker, T. Barnes, & J. Beck (Eds.), *Proceedings of the1st International Conference on Educational Data Mining, EDM 2008* (pp. 117–126). Montreal, Canada.

Shih, B., Koedinger, K. R., & Scheines, R. (2010). Unsupervised discovery of student learning tactics. In R. S. J. D. Baker, A. Merceron, & P. I. Pavlik (Eds.), *Proceedings of the 3rd International Conference on Educational Data Mining, EDM 2010* (pp. 201–210). Worcester: International Educational Data Mining Society.

Stamper, J., Barnes, T., & Croy, M. (2011). Enhancing the automatic generation of hints with expert seeding. *International Journal of Artificial Intelligence in Education, 21*(1–2), 153–167. doi:10.3233/JAI-2011-021.

Tai, M., Arroyo, I., & Woolf, B. (2013). Teammate relationships improve help-seeking behavior in an intelligent tutoring system. In H. C. Lane, K. Yacef, J. Mostow, & P. Pavlik (Eds.), *Lecture Notes in Computer Science: Artificial Intelligence in Education* (Vol. 7926, pp. 239–248). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-39112-5_25.

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

VanLehn, K., Lynch, C., Schulze, K., Shapiro, J. A., Shelby, R., Taylor, L., & Wintersgill, M. (2005). The Andes physics tutoring system: lessons learned. *International Journal of Artificial Intelligence in Education, 15*(1), 147–204.

Vassileva, J., McCalla, G.I., & Greer, J.E. (2016). From small seeds grow fruitful trees: How the PHhelpS peer help system stimulated a diverse and innovative research agenda over 15 years. *International Journal of Artificial Intelligence in Education, 26*(1).

Walker, E., Rummel, N., & Koedinger, K. R. (2014). Adaptive intelligent support to improve peer tutoring in algebra. *International Journal of Artificial Intelligence in Education, 24*(1), 33–61. doi:10.1007/s40593-013-0001-9.

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

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in educational theory and practice* (pp. 279–306). Hillsdale: Erlbaum.

Wittwer, J., & Renkl, A. (2008). Why instructional explanations often do not work: a framework for understanding the effectiveness of instructional explanations. *Educational Psychologist, 43*(1), 49–64. doi:10.1080/00461520701756420.

Wood, H., & Wood, D. (1999). Help seeking, learning and contingent tutoring. *Computers & Education, 33*(2/3), 153–169.

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). San Diego: Academic.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. *American Educational Research Journal, 45*(1), 166–183. doi:10.3102/0002831207312909.