INSTRUCTION BASED ON ADAPTIVE LEARNING TECHNOLOGIES

Vincent Aleven, Elizabeth A. McLaughlin, R. Amos Glenn, Kenneth R. Koedinger

Human-Computer Interaction Institute

Carnegie Mellon University

aleven@cs.cmu.edu, mimim@cs.cmu.edu, amosg@cs.cmu.edu, koedinger@cmu.edu

Draft. Please do not distribute.

INSTRUCTION BASED ON ADAPTIVE LEARNING TECHNOLOGIES

Introduction

A common intuition is that instruction is most effective if it takes into account (a) that learners are different, and (b) that they change as they learn. But learners differ in a great many ways, for example in their knowledge state, interest, goals, affective state, strategic behaviors, and learning styles. To what learner differences should instruction adapt? Also, how can adaptive instruction take into account that learners change continuously? Which ways of adapting are most effective? An alternative view is that learning technologies should not only adapt to learner differences, but also to learner similarities. How can a deep understanding of the demands a task domain makes on *all* learners be used both to adapt system design to common rough spots?

The idea of adapting to learners to individualize instruction is not new. For example, mastery learning is a well-known and effective way of adapting instruction to individual students' knowledge growth (Bloom, 1968; Keller, 1968; Kulik, Kulik, & Bangert-Drowns, 1990). In spite of this success, mastery learning is not implemented on a large scale in regular instructional practice. Adaptive forms of instruction, such as mastery learning, can encounter barriers to adoption including resource demands (e.g., frequent formative assessment and individualized remedial practice or instruction), practical challenges for educators (e.g., how to manage a student with 3rd grade reading needs and 2nd grade mathematics needs who is, by age and socially, in 4th grade), and a perception of unfairness because different students are assigned different work to accomplish the same milestones.

Technology can help address some of these barriers. By now, advanced learning technologies such as intelligent tutoring systems can support many forms of adaptivity and individualization, in ways that would be difficult to manage for teachers (Koedinger et al., 2013).

They can assess learners along many psychological dimensions, including knowledge, affect, and metacognition (Aleven & Koedinger, 2013; Conati & Kardan, 2013; Sottilare, Graesser, Hu, & Holden, 2013). They can adjust their pedagogical decision making accordingly. They often do so without separate assessment activities that take away from learning time because these systems can assess students as they work on instructional activities. For example, a "cognitive mastery" approach implemented in intelligent tutoring software has been shown to substantially enhance student learning (Corbett, McLaughlin, & Scarpinatto, 2000). It has also turned out to be a viable and practical method in schools, addressing some of the practical obstacles noted. It is implemented in Cognitive Tutor software used by hundreds of thousands of middle-school and high-school students annually (Koedinger & Corbett, 2006).

Adaptive instruction builds on a variety of theoretical perspectives, including work documenting aptitude-treatment interactions (ATIs; e.g., Cronbach & Snow, 1977), individual differences in learning (Jonassen & Grabowski, 1993), expertise reversal (e.g., Kalyuga 2007; Kalyuga, Ayres, Chandler, & Sweller, 2003), the Zone of Proximal Development (Vygotsky, 1978), the model-scaffold-fade paradigm (Collins, Brown, & Newman, 1989), and the Assistance Dilemma (Koedinger & Aleven, 2007; Koedinger, Pavlik, McLaren, & Aleven, 2008). For example, educational research on ATIs (Cronbach & Snow, 1977; Snow, 1989; Kalyuga et al., 2003) has documented many instances of where a choice of an effective instructional treatment depends on learner characteristics such as prior knowledge.

In this chapter, we take stock of the state of empirical research regarding the value of adapting instruction to the demands of the domain, learner characteristics, and the learner's path in the ongoing learning activity. Our review includes work on advanced learning technologies, such as intelligent tutoring systems, conversational agents, and educational games. However, the

focus is on empirical work and not on technology. In this sense, our chapter is different from previous reviews of adaptivity in learning technologies (e.g., Brusilovsky, 2001; Vandewaetere & Clarebout, 2014; Vandewaetere, Desmet & Clarebout, 2011; VanLehn, 2006; 2016), which focus on technology aspects. Our chapter is a selective review. For the topics discussed, we searched broadly for articles that compared adaptive and non-adaptive instruction, as well as for articles that tested whether interactions might exist between instructional treatments and learner characteristics.

What Does It Mean To Be Adaptive?

We define *adaptivity* as follows: A learning environment is adaptive to the degree that (a) its design is based on data about common learner challenges in the target subject matter, (b) its pedagogical decision making changes based on psychological measures of individual learners, and (c) it interactively responds to learner actions (cf. Aleven et al., 2015; Aleven, Beal, & Graesser, 2013). According to this definition, some systems may be more adaptive than others (i.e., adaptivity is a matter of degree, not a binary property).

For all parts of the definition, adaptivity requires data about learners. Part a of the definition captures the design of systems based on data from cognitive task analysis, such as qualitative data from interviews (Clark et al., 2007) or think alouds (Ericsson & Simon, 1984) as well as quantitative data from student performance on tasks given in experiments on paper (e.g., Koedinger & Nathan, 2004) or as part of an existing educational technology (e.g., Stamper & Koedinger, 2011). Parts b and c of the definition capture ways in which the running system adjusts its behavior based on data it gathers about each student, as students use it. The system can adapt to students over a short time span, in reaction to a single student action, or over a long time span, in reaction to a student state or trait identified over many student actions.

In addition, according to part a of our definition, a system may be adaptive at *design time* even if it is not adaptive at *run time* (i.e., as it is being used by students), perhaps extending how the term "adaptive" is commonly used. A system is adaptive at design time if it is designed in a way that is responsive to the learning demands that the domain produces that are largely the same for many learners (e.g., challenges or hurdles that are the same across learners). For example, a video lecture might be seen as non-adaptive form of instruction. The video is the same for all learners and in all situations. It does not satisfy parts b and c of our definition. Nonetheless, a video lecture could be viewed as adaptive if it has been designed based on a careful analysis of data about student learning in the given domain (part a of the definition). This would be adaptivity not to individual learners, but to the demands that the task domain makes on learners in general. The video would be in instance of design-time adaptivity.

Adaptive to What?

What should instruction adapt to in order to be more effective than one-size-fits-all instruction? There are many learner characteristics to consider, by one accounting as many as 30 (Jonassen & Grabowski, 1993). Which characteristics are most worth adapting to is an empirical question. In our review of the empirical literature on adaptive learning technologies, we distinguish five broad groups of learner characteristics, shown as the rows of the Adaptivity Grid in Table 1. A key characteristic is student knowledge. There is ample evidence in various literatures (e.g., Corbett et al., 2000; Dochy et al., 2002; Jonassen & Grabowski, 1993; Kalyuga, 2007; Tobias, 1994) that students differ significantly in their prior knowledge related to given subject matter and that the effectiveness of instructional treatments interacts with students' knowledge. A key idea is further that instruction needs to be designed with the knowledge demands of a domain clearly in mind, and that uncovering these knowledge demands is best done using data. As a

Table 1. The Adaptivity Grid: columns indicate the different time scales of adaptation (i.e., what to adapt), rows indicate learner characteristics for which instruction is adapted (i.e., what to adapt to), and each cell provides empirical evidence (i.e., relevant studies reported in this chapter).

	Design Loop	Task Loop	Step Loop
Prior knowledge and knowledge growth	Koedinger & McLaughlin, 2010; Koedinger, Stamper, McLaughlin, & Nixon, 2013b; Lovett, Meyer, & Thille, 2008; Stamper & Koedinger, 2011	Arroyo, Mehranian, & Woolf, 2010; Atkinson, 1972; Beck, Woolf, & Beal, 2000; Corbett, McLaughlin, & Scarpinatto, 2000; Kalyuga & Sweller, 2004, 2005; Muldner & Conati, 2007; Nückles, Hübner, Dümer, & Renkl, 2010; Oksa, Kalyuga, & Chandler, 2010; Salden, Aleven, Schwonke, & Renkl, 2010; Salden, Paas, & van Merriënboer, 2006b; Van Merriënboer et al., 2004	Conati & VanLehn; 2000; Mitrovic, Ohlsson, & Barrow, 2013
Students' path through a problem: Strategies, errors	Adams et al., 2014; Booth, Lange, Koedinger, & Newton, 2013; Koedinger & Anderson, 1998; McKendree, 1990; McLaren et al., 2016	Anderson et al., 1995; McLaren et al., 2012; Roll, Aleven, & Koedinger, 2010	Anderson, Conrad, & Corbett, 1989; Chi, VanLehn, Litman, & Jordan, 2011; Lee, Rowe, Mott, & Lester, 2014; Rittle-Johnson & Star, 2007; Rowe & Lester, 2015; Stamper, Eagle, Barnes, & Croy, 2013; Waalkens, Aleven, & Taatgen, 2013
Affect, motivation	D'Mello, Lehman, Pekrun, & Graesser, 2014; Lehman, D'Mello, Strain, Mills, Gross, Dobbins, . . . Graesser, 2013	Anand & Ross, 1987; Baker et al, 2013; Baker et al., 2009; Baker et al., 2008; Bernacki & Walkington, 2014; Heilman et al., 2010; Walkington, 2013; Walkington & Bernacki, 2015; Walkington & Sherman, 2012	D'Mello et al., 2010; D'Mello, Olney, Williams, & Hayes, 2012; Forbes-Riley & Litman, 2011
Self-regulated learning strategies, metacognition, effort	Aleven & Koedinger, 2002	Arroyo et al., 2007; Long, 2015; Long & Aleven, 2013a; Long & Aleven, 2013b; Mitrovic & Martin, 2007	Aleven, McLaren, Roll, & Koedinger, 2010; Duffy & Azevedo, 2015; Kinnebrew, Szegedy, & Biswas, 2014; Leelawong & Biswas, 2008; Mathan & Koedinger, 2005; Roll, Aleven, McLaren, & Koedinger, 2011; Wagster, Tan, Wu, Biswas, & Schwartz, 2007
Learning styles		Brown, Brailsford, Fischer, & Moore, 2009; Constantinidou & Baker, 2002; Cook, Thompson, Thomas, & Thomas, 2009; Ford & Chen, 2001; Graf & Kinshuk, 2007; Mampadi, Chen, Ghinea, & Chen, 2011; Massa & Mayer, 2006; Popescu, 2009, 2010; Tseng, Chu, Hwang, & Tsai, 2008;	

Draft. Please do not distribute.

second group of characteristics, a system may respond to the specifics of a student's path through a learning activity, including a student's solution strategy, specific errors, requests for help and assistance, and other elements of the problem state. Third, researchers in advanced learning technologies have become very interested in how instruction can assess learners' affective or motivational state, and how instruction might adjust to these characteristics on the fly. Fourth, we consider to what degree it has been shown to be fruitful for instruction to be adaptive to learners' self-regulatory processes, including metacognitive processes. Finally, it is often thought that instruction should adjust to students' learning styles.

In order to adapt to individual differences in the task-loop or step-loop, the system needs to assess these characteristics for individual students, especially for characteristics that can reasonably be expected to change over the course of instruction (e.g., knowledge of the targeted learning objectives, affective or motivational state, and so forth). The fields of AI in Education and Educational Data Mining have produced many techniques for inferring student characteristics from available data, under the banner of student modeling (e.g., Desmarais & Baker, 2012). This topic however is outside the scope of the current chapter.

How and When to Adapt?

How and when should instruction adapt or be adapted to learner similarities and differences? In principle, any instructional feature could be adapted, based on learner characteristics. We organize the many ways in which instruction can vary and change into three broad categories, with different time scales, from slowest to fastest: Design-loop adaptivity, task-loop adaptivity, and step-loop adaptivity, shown as the columns of the Adaptivity Grid (see Table 1).

Design-loop adaptivity involves data-driven decisions made by course designers before

Draft. Please do not distribute.

and between iterations of system design, in which a course or system is updated based on data about student learning, specifically, data collected with the same system or course. As a result, the system becomes more adaptive (or adapted) to the demands of the task domain. This corresponds to part (a) of our definition of adaptivity, above. In design-loop adaptivity, instruction is adapted to similarities among learners.

Task-loop adaptivity involves data-driven decisions the system makes to select instructional tasks for the learner. In task-loop adaptivity, the instruction is changed one task at a time; a paradigmatic example is individualized task selection, for example as it occurs in cognitive mastery approaches (Corbett & Anderson, 1995; Corbett et al., 2000).

Step-loop adaptivity involves data-driven decisions the system makes in response to individual actions a student takes within an instructional task. In step-loop adaptivity, instructional features that operate *within a task or learning activity* are changed based on learner characteristics. The distinction between step-loop functionality and task-loop functionality is often made in theoretical accounts of advanced learning technologies (e.g., VanLehn, 2006, 2011, 2016), as a practical way of thinking about system behavior.

Nature of Empirical Evidence Studied

The two dimensions of adaptivity (i.e., when/how to adapt, and what to adapt to) define a design space for adaptive instructional systems, shown in the Adaptivity Grid of Table 1. Each cell represents a class of adaptive mechanisms, typically consisting of a method for assessing student characteristics and one for adjusting the instruction based on this assessment. In the remainder of the chapter, we discuss empirical evidence regarding the Adaptivity Grid. For each cell, we look at empirical work that tested the effect of the particular class of adaptive mechanisms captured in that cell. We look at experimental studies that compared outcomes

resulting from adaptive versus non-adaptive versions of the same instruction. We also look at studies that tested – in the context of advanced learning technologies – whether interactions exist between student characteristics and instructional treatments. The value of adaptivity is demonstrated either when an adaptive instructional treatment produces more favorable outcomes than corresponding non-adaptive instruction, or when a crossover interaction is found between a choice of instructional treatment and a learner characteristic, such that learners with the characteristic do better with one instructional treatment, but learners without the characteristic do better with a different instructional treatment.

Adapting to Prior Knowledge and Knowledge Growth

In technology-driven courses, as in traditional courses and classrooms, students enter with varying degrees of knowledge of and experience with the targeted domain. A one-lesson-fits-all approach to learning and instruction becomes implausible when learning environments are filled with such a diversity of incoming knowledge. The heterogeneity of knowledge (i.e., variance in incoming, on-going and learning rates) among learners is considered by many to be the most influential factor in learning (Dochy, Segers, & Pletinckx, 2002; Tobias, 1994). Thus, a key goal of adaptivity in learning environments is to improve methods for assessing prior knowledge and knowledge growth, and then adapting the instruction accordingly. These adaptations can be done in the design-loop, task-loop, or step-loop as can be seen in Table 1. All three ways of adapting to prior knowledge have been shown to be effective in enhancing desirable outcomes of instruction.

Design-Loop Adaptations To Student Knowledge

Design-loop adaptation is aptly demonstrated with the redesign of a course or tutoring

system that occurs as a result of discovering deficiencies in the knowledge component model that underlies the instruction. Typically, this discovery process involves student data, especially student data from the same course. A knowledge component model is a fine-grained decomposition of the knowledge targeted in the instruction, together with a mapping that specifies which activities involve which knowledge components (Koedinger, Corbett, & Perfetti, 2012; Stamper & Koedinger, 2011). Knowledge component models are commonly used in the design and analysis of intelligent tutoring systems but apply equally to online courses or instruction without technology.

Analysis of log data, in which activities are tagged based on a knowledge component model, may reveal *hidden skills*, that is, knowledge within the instructional objectives that was missed by the instructor or course designers and therefore is missing from the knowledge component model. It may also reveal missed generalizations that students make, or assumed generalizations that students miss, meaning that the knowledge component model captures knowledge at the wrong level of abstraction. These kinds of deficiencies can be quite difficult to avoid in the original knowledge component model underlying a course or tutoring system, but can adversely affect the instruction's efficiency or effectiveness.

The redesign of a blended statistics course (Lovett, Meyer, & Thille, 2008) illustrates that design-loop adaptivity can dramatically enhance the effectiveness and efficiency of instruction. Using course log data and pre/post test data, knowledge component analysis uncovered deficiencies in the knowledge component model underlying the course, including missed generalizations and incorrectly assumed generalizations. Based on this analysis, the course objectives were revised, new activities were added and others were deleted to ensure all objectives received the right amount of practice. The redesigned course led to dramatically better

learning outcomes in half the time, compared to the original course (i.e., the course without the new design-loop adaptations).

As a second illustration, Stamper and Koedinger (2011) redesigned a unit of a geometry tutor after discovering specific knowledge components were missing from the tutor's knowledge component model. Using visualization tools (e.g., learning curves) and similar data-mining techniques as were used in the Lovett et al. (2008) study, they discovered a latent planning skill (i.e., knowing to decompose) that was hindering students from successfully computing the area of multi-shaped figures. They redesigned the tutor in a number of ways to address these new skills (e.g., they created new problems targeting the planning skill, included more practice problems with the targeted skill, and added hints to help students apply the planning skill). They found that students using the adapted tutor had more efficient learning (i.e., less time to mastery) and higher performance on a posttest, compared to the original tutor prior to the re-design (i.e., the tutor without the design-loop adaptivity).

In a third example of design-loop adaptation, Koedinger and McLaughlin (2010) based a system redesign on data showing that algebra students are worse at writing expressions (e.g., 800-40x) for two-operator story problems (e.g., "Anne is 800 meters from the dock rowing at 40 meters per minute ...") than they are at writing expressions (e.g., 40x and 800-y) for two otherwise equivalent one-operator story problems (Heffernan & Koedinger, 1997). They hypothesized that practice on substitution problems (e.g., "Substitute 40m for x in 800 – x") would address the gap. They added and indeed found better learning in comparison to a control that involved practice on simple one-step story problems.

These examples illustrate that it can be effective to adapt courses or tutoring systems to the specific knowledge demands of a given domain, even without run-time adaptivity. As

illustrated, a knowledge component approach can be used to uncover these demands by finding common patterns in student data (Aleven & Koedinger, 2013; Stamper & Koedinger, 2011), refine learner models, and adapt the corresponding instruction to learner *similarities.* A major benefit to a knowledge-component approach, as distinct from other learner modeling (e.g., ALEKs selects next problem based on Knowledge Space Theory; Falmagne, Albert, Doble, Eppstein & Hu, 2013; Falmagne et al., 1990) is the development of explanatory models that can easily be interpreted, and therefore used to facilitate the redesign of new tasks adapted to student needs (Koedinger, Stamper, McLaughlin, & Nixon, 2013).

Task-Loop Adaptations To Student Knowledge

There is good evidence that task selection based on assessment of individual students' knowledge state can substantially contribute to the effectiveness of instruction. For example, more than 40 years ago, Atkinson (1972) studied task selection techniques for learning German vocabulary. Of the four approaches he evaluated, the two that used student response history for automated item selection showed the greatest gain on a delayed posttest, more than both the learner control and random selection conditions.

Similar results have been obtained with intelligent tutoring systems developed using cognitive theory and learner modeling principles. For example, Cognitive Tutors, a type of intelligent tutoring system grounded in cognitive theory (Anderson et al., 1995), implement a variation of mastery learning, called cognitive mastery task selection. In this approach, the system dynamically assesses student knowledge using a Bayesian model (called Bayesian Knowledge Tracing; Corbett & Anderson, 1995). Based on this assessment, the system selects tasks that exercise knowledge components that a given student is not likely to have mastered. This approach to adaptive task selection, as used in an intelligent tutoring system, was shown to

substantially improve the effectiveness and efficiency of student learning, compared to giving all students the same problem set (Corbett, McLaughlin, & Scarpinatto, 2000).   It has become a standard in commercially-available tutoring systems (Cognitive Tutors).

A different form of task-loop adaptivity to student knowledge, namely, adaptive selection and presentation of worked examples to students, has also been shown to be effective. Muldner and Conati (2007) devised a method to adaptively select the best example to present to a student to promote learning by analogical problem solving (in which the student solves a problem aided by having an analogous example at hand). The adaptive method selected examples based on students' domain knowledge, their tendencies for using certain strategies for learning from examples known to be effective, and the similarity between a problem and candidate example. Adaptively presented examples led to greater use of self-explanation strategies known to be conducive to learning, compared to non-adaptively selected examples, but there was no difference in learning gains. On the other hand, Salden, Aleven, Schwonke, and Renkl (2010), in the context of an intelligent tutoring system for high-school geometry, found that adaptively transitioning from worked examples to tutored problem solving has a greater impact on learning than fixed fading of examples or no worked examples (i.e., tutored problem solving by itself). Here, the decision made in the task loop was which steps in the next problem should be presented to the student as worked-out steps to be explained, and which steps should be open, for the student to solve. This decision was based on the system's assessment of how well students explained the worked examples, using the Bayesian Knowledge Tracing method described above (Corbett & Anderson, 1995).

The extensive line of work on expertise reversal, which has provided abundant evidence that instruction for novice learners should be different than that for advanced knowledge learners

both in well-defined domains (e.g., algebra- Kalyuga et al., 2003; Kalyuga, 2007) and ill-defined domains (e.g., literary text: Oksa, Kalyuga & Chandler, 2010; journal writing: Nückles, Hübner, Dümer & Renkl, 2010; air traffic control: Salden, Paas & van Merrienboer, 2006a, 2006b), has also demonstrated that adapting to student knowledge in the task loop can yield more effective instruction. In a further development of expertise reversal, Kalyuga and Sweller (2004) tested adaptation to student knowledge based on a (domain-specific) rapid dynamic assessment method (RDA). This assessment method was based on the assumption that advanced learners choose a different solution path than novice learners (i.e., the more knowledgeable, the more likely a learner will skip earlier, easier steps and jump to later steps). In four separate experiments, Kalyuga and Sweller (2004) demonstrated the effectiveness of the RDA method for initial training placement based on prior knowledge (e.g., fully worked examples vs. faded examples) and adapting instruction based on knowledge growth in real time.

Although assessing cognition is fundamental to adaptation, it is hard to isolate knowledge growth from load and effort. Kalyuga and Sweller (2005) added a cognitive efficiency rating (using self-reported measures of cognitive load with performance measures) and found greater knowledge and cognitive efficiency gains. Van Merrienboer et al. (2004) used a four-component instructional design model (4C/ID-Model) for task selection that accounts for cognitive load and mental effort when assessing expertise. Arroyo, Mehranian, and Woolf (2010) demonstrated the promise of a task selection policy that combines assessment of mastery with estimates of effort. In other projects, discussed below, aspects of student affect are taken into account, in combination with cognitive factors, in adaptive task selection policies.

Step-Loop Adaptations to Student Knowledge

Adaptation to prior knowledge and knowledge growth has also been shown to be

effective in the system's step loop. For example, Conati and VanLehn (2000) and, later, Hausmann, Nokes, VanLehn, and Gershman (2009), studied effects of adaptive support for self-explanations of worked-example steps. They implemented a Self-Explanation Coach (SE-Coach) in a tutor for college-level physics problem solving (Conati, 2013; 2016; Conati & VanLehn, 2000). The SE-Coach adaptively selected the steps of worked examples, based on the tutor's assessment of student understanding, captured in the system's learner model. It also provided a structured template interface for two types of self-explanations, as well as feedback on explanations entered in this interface. The adaptive support for self-explanation led to greater learning gains, compared to a control conditions that was given prompts to explain but no adaptive support for self-explanation. This effect was found only for students with low prior knowledge. Thus, similar to results found in Kalyuga and Sweller's (2005) task-loop adaptation, as students became more knowledgeable, the less they needed structured, adaptive help and the more likely scaffolding interfered.

Mitrovic, Ohlsson, and Barrow (2013), investigated effects of adaptive positive feedback on student solutions with SQL-Tutor, a tutoring system that helps students learn to write computer code to query a database. The type of tutor used (called a constraint-based tutor) typically provides error feedback only (Mitrovic & Ohlsson, 2006; Ohlsson, 2016). The authors hypothesized that positive feedback to a student's attempted solution, given in addition to error feedback, would help reduce student uncertainty and thereby help students learn more effectively or efficiently. They devised an adaptive mechanism that gave positive feedback on solution aspects that were correct, but about which the student might be uncertain. Potential uncertainty was identified based on the submitted solution, the student's knowledge (captured in the system's long-term student model) and the state of his or her interaction with the system. In a

classroom experiment, Mitrovic et al. (2013) compared the revised tutor, which gave adaptive

feedback on successes and errors, with the original tutor, which gave feedback on errors only,

and found the students in the experimental condition reached mastery in half the time of the

control condition.   This approach is closely related to that by Forbes-Riley and Litmann (2011),

described below.

Discussion of Adapting to Student Knowledge

Much evidence points to the benefits of adapting instruction to student knowledge

congruent with the notion that prior knowledge is a key influence on student learning. In design-

loop adaptation, there have been some striking demonstrations of improved effectiveness due to

offline analysis in which educational data mining methods provided insight into the demands of

the domain, led to refined knowledge component models, and informed redesign of the course or

tutoring system. This line of work leads one to consider instructional design as an iterative

process in which data from an earlier system version is key in creating the next version, which is

more adaptive to *all* learners. Further, task-loop adaptivity to students' knowledge growth may

well be the single cell in our Adaptivity Grid where adaptive instruction is most effective.

Adaptive forms of task selection based on dynamic assessment of students' evolving knowledge

were shown to be more effective to non-adaptive instruction that presented the same sequence to

all students (e.g., the work on cognitive mastery by Corbett et al., 2000). Although this work has

been impactful, it would be useful to see one or more replications of these results. The evidence

is not as prolific regarding the value of adapting to knowledge growth at the step level, although

here too we see some interesting demonstrations. There may be room for interesting innovations

in the step loop and the task loop. For example, adaptively fading support may be effective (e.g.,

VanLehn et al., 2000; see also Collins et al., 1989; Salden et al., 2010). The extensive line of

studies focused on the expertise reversal effect might prove fertile ground for the design of adaptive instruction. Finally, more and more, researchers are developing task selection methods that combine a range of learner factors, rather than focusing on student knowledge by itself (e.g., Arroyo et al., 2014; Mazziotti et al., 2015; Grawemeyer, Mavrikis, Holmes, & Gutierrez-Santos, 2015).

Adapting to Student Strategies or Errors

Adapting to student strategies and errors means using data about the character and frequency of student strategy use and error patterns to make decisions about (select or change) some element of instruction. Such adaptations have been tried and were found effective in all three loops (i.e., in the design, task, and step loops).

Design-Loop Adaptations To Strategies or Errors

Educational technology design can often be enhanced by adapting it to data-identified similarities in students' strategies or errors. An example of adaptation to *non-obvious* similarities in student strategies comes from cognitive task analysis toward design of the Algebra Cognitive Tutor (Koedinger & Nathan, 2004). It was discovered that beginning algebra students used informal strategies to perform better on algebra story and word problems (66% and 62% correct, respectively) than on matched equations (43% correct), rather than the normative strategy of translating a story into an equation and solving the equation. The informal strategies included iteratively generating an estimate for the unknown value, following the verbal specification of the computations (thus not needing an equation), and testing whether the given result is achieved. A design loop adaptation was implemented to capitalize on these findings. Although an initial Algebra Cognitive Tutor unit design followed an existing textbook approach of scaffolding

students by prompting them to first write an equation before answering problem-solving

questions, the data suggested doing so was not good cognitive load management because it

required students do something harder (equation symbolization) before something easier (story

problem solving). The adaptive design idea was to switch this order, that is, to have students do

something easier first (solve the problem working out the computations needed) to bridge to

something harder (generalize and express the computations in algebraic symbols). A random

assignment experiment reported in Koedinger and Anderson (1998) demonstrated that this

adaptive design idea does, indeed, produce better student learning, compared to the original

design that followed the textbook order (i.e., prompt for equations first).

     In addition to making design loop adaptations in reaction to student *strategy* data, design

loop adaptations can also be made in reaction to data on similarities in student *errors*. In

research on the Geometry Proof Tutor, McKendree (1990) identified certain strategic proof

planning skills that were particularly difficult for students. She adapted the design of the tutor so

that the tutor's error feedback (given when students are stuck on a step) provided instruction on

these strategies (e.g., by pointing out the goal on which to focus next in the proof). She

demonstrated, in a random assignment experiment, that a tutor with these error messages

produced better student learning than the original tutor, which provided correctness feedback

only (i.e., only informed the student whether each proof step was correct or not). Note, this

adaption, changing the strategic focus on all error feedback messages, was applied across all

students (in the treatment condition) based on a similarity across students (they all tend to have

difficulty with strategic planning decisions). Thus, it is a design loop adaption.

     As another example of design loop adaptations, a number of researchers have used

domain-specific data or literature on common student errors to design tutoring systems that

present incorrect or erroneous worked examples to students. They performed random assignment experiments to investigate whether the addition of incorrect worked examples enhances student learning relative to comparable controls, namely, the same tutoring systems without incorrect worked examples but with correct examples and/or problem-solving practice. Booth, Lange, Koedinger, and Newton (2013) created alternative versions of an algebra equation-solving tutor that included, respectively, only correct examples, only incorrect examples, or both. They found that incorrect examples enhanced student learning, compared to the other tutor versions. Their design-loop adaptation is based on substantial data collection and analysis of persistent and problematic errors in Algebra I (cf., Booth, Barbieri, Eyer, & Pare-Blagoev, 2014). Similarly, two studies (Adams et al, 2014; McLaren et al., 2016) with a math tutoring system for learning decimals, called AdaptErrEx, found enhanced learning in a condition with incorrect examples as measured on a delayed post-test, as compared to tutored problem solving without examples. Their design-loop adaptation was based on substantial data and analysis that identified students' most common decimal misconceptions. Isotani, McLaren, and Altman (2010) describe how they adapted instruction to this data on common student errors by designing incorrect examples that targeted the identified misconceptions. Thus we see that redesigning systems based on data about student strategies and errors in the given domain can lead to re-designed instruction that is more effective than the original, even if the system does not adapt dynamically to student differences.

Task-Loop Adaptations To Strategies or Errors

The AdaptErrEx system (Goguadze, Sosnovsky, Isotani, & McLaren, 2011; McLaren et al., 2012) provides an example of how task-loop adaptations to differences in student errors can be implemented. This system adapts the choice of erroneous examples to present to an

individual student, based on pre-test differences in student error patterns and associated

misconceptions. If pre-test errors suggest the student may believe that the bigger the whole

number to the right of the decimal, the bigger the decimal, then a good choice of erroneous

example might be: "0.25 is bigger than 0.5" because this student is likely to make that same

error. Conversely, if pre-test errors of a student suggest she has the misconception that more

digits means a smaller decimal, now "0.5 bigger than 0.75?" becomes a good adaptive choice to

help the student debug her misconception. As mentioned, erroneous examples helped students

learn better, although the value of adaptive selection was not confirmed.

As another instance of adapting to errors in the system's task loop, the Invention Lab

(Roll, Aleven, & Koedinger, 2010) selects new contrasting cases based on the specific

shortcomings in students' attempts at inventing procedures that achieve desirable goals (e.g.,

develop a measure of variability in a numeric variable sample such as how high an object

bounces when dropped on a trampoline). Cases can be considered to be tasks within a

curriculum or project and so we place the Invention Lab in this section (and also because of its

similarity with AdaptErrEx). However, cases might, instead, be considered as steps in an

invention problem/activity and then InventionLab is an instance of step-loop

adaptation. Experimental results suggest that having students attempt to invent procedures and

adaptively get new cases based on limitations in the proposed procedure yields better transfer

than having students evaluate given procedures, although the experiment did not try to isolate the

value of adaptivity.

Beck, Woolf and Beal (2000) investigated whether statistical and machine learning

techniques could be used for effective task-loop adaptivity. Using linear regression and

reinforcement learning, they created an adaptive agent, called ADVISOR, that adaptively

selected the next topic, problem, or hint based on the student's level of prior proficiency and

level of cognitive development, the difficulty of the current topic and the current problem, and

the student's prior attempts at answering, including hints seen. The objective of ADVISOR's

adaptive policy was to minimize the time the student spends per problem. (The choice of

objective was driven in part by the desire to have a proof-of-principle demonstration for the

given approach to using machine learning.) Embedded in an intelligent tutor for arithmetic

(AnimalWatch), ADVISOR led to a significant reduction in the time spent per problem

compared to using AnimalWatch as-is. This project demonstrated that a machine-learned

adaptive policy for tutoring can influence students' learning processes, although learning

outcomes were not measured.

Step-Loop Adaptations To Student Strategies or Errors

Error feedback and next-step hints are forms of step-loop adaption that are common in

intelligent tutoring systems (cf., Koedinger et al., 2013a; VanLehn, 2006; 2011). This kind of

adaptation gives different students different instructional feedback and hints (i.e., it adapts

instruction to student needs) depending on the specific errors those students make or the specific

strategies they pursue. Dynamic adaption to specific student errors requires a system that can

monitor student step-by-step actions in the given computer interface and evaluate those actions

as correct or incorrect or, in less well-defined domains, as more desirable or less desirable. Some

forms of intelligent tutoring systems can not only evaluate the correctness of student solutions

and actions, but can generate correct solutions to problems and complete partially complete

solutions. Such tutors can provide another form of adaptation, namely, hints as to what to do

next. If the problem space is such that multiple solution strategies are possible within a given

problem, then dynamic adaptation of hints (so that the next-step hint is adaptive to student

strategies), requires that the tutoring system is capable of recognizing of generating alternative solutions. As a simple example: If one student is adding fractions with unlike denominators, like ¼ + ⅙, and had entered 24 as the converted denominator for ¼, she will get a different next-step hint than a student who had entered 12 as the converted denominator of ¼. As another example, this time from the domain of introductory computer programming, when stuck in a programming problem that requires repeated steps, a student pursuing a looping solution will get a different next step hint (e.g., toward entering terminating condition on a while loop) than a student pursuing a recursive solution (e.g., a hint toward entering a base condition to exit the recursion).

A number of experiments have explored variations on whether, how, and when giving solution state feedback or next-step hints aid learning (see Aleven, McLaren, Roll, & Koedinger, 2016; Koedinger & Aleven, 2007, for a review of some of these studies). One of the most powerful early results comes from experiments with the LISP programming tutor (Anderson, Conrad, & Corbett, 1989) where it was found that students who received step-by-step error feedback and as-needed next-step hints learned more than students with a typical set of problems and program solution evaluation only at the end of each problem, and they did so in one third the time.

Work by Stamper, Eagle, Barnes, and Croy (2013) showed the value of next-step hints that adapt to the solution path that the student is following. They used data from past student solutions to problems, with machine learning methods, to automatically generate next-step hints for students using an online environment for practicing logic proofs. The original version of the online proof system allowed any logically correct step but gave no suggestions or feedback as to which of the many possible correct next steps would advance the proof. The machine-learned hints suggested a next step that (in past student work) was a frequent choice on the way to a

completed proof, from the given state. In a quasi-experiment, the adaptive next-step hints helped

students persevere, do better in the tutor, and obtain better course grades, compared to a the

original online proof system, which provided no suggestions for what to do next. More generally,

our recent review of the literature on the value of next-step hints indicates that these hints help

learning, albeit to a limited degree (Aleven et al., 2016).

Potential benefits of adapting to students' strategy choices within the step loop were

explored by Waalkens, Aleven, and Taatgen (2013). They compared learning with three versions

of a tutoring system for equation solving that differed only in the range of student strategies that

the system recognized as correct and for which it provided tutoring. One system version

recognized and supported only a standard strategy, determined based on cognitive task analysis

and review of textbooks. Another version dealt with this standard strategy as well as minor

variations. A third recognized major and minor strategy variations. In the last two versions,

students were free to use whichever strategy variation they preferred, within the range

recognized by the system. Surprisingly, Waalkens et al. found no difference in learning or

enjoyment due to these different degrees of adaptivity (or strategy freedom). These results are in

line with the literature on supporting strategic flexibility in algebra (e.g., Rittle-Johnson & Star,

2007). They suggest that in order to help students learn to use multiple strategies and acquire

strategic flexibility, more is needed than merely being able to follow along with student strategy

choices.   For example, when instructional goals include use of multiple alternative domain

strategies, these strategies should be practiced separately. When the instructional goals include

strategic flexibility, additional activities may need to be designed in which students have the

freedom to select their own strategies (as in the study with Lynnette), perhaps with added support

for reflecting on the choice of strategies or comparing different solutions. It may be helpful as

well as to provide activities that provide focused practice with strategy choice (cf. Rittle-Johnson & Star, 2007).

Increasingly, work focuses on adaptively selecting the next tutor action or dialogue move in the system's step loop; often, machine learning is used to create policies are adaptive to student errors and strategy choice, and often to other variables as well. Early work by Beck et al. (2000) was mentioned above. As another example, in work with a tutorial dialogue system, Chi, VanLehn, Litman, and Jordan (2011) used reinforcement learning (a particular machine learning technique) to generate a policy that selects between *tell* and *elicit* moves by the tutor agent (i.e., whether to provide information to the student or ask the student to generate it, a key decision in tutorial dialogue). In a lab study, this policy was shown to lead to better learning than a policy that was deliberately "counter-adaptive" (i.e., selected *tell* when *elicit* was called for, and vice versa), demonstrating that choice between *tell* and *elicit* influences student learning. That left open the question, however, of whether this adaptive policy leads to better student learning, compared to a reasonable non-adaptive alternative. Similarly, Murray, VanLehn and Mostow (2004) used a Dynamic Bayesian Network to select the next tutor move in tutoring systems for reading and for calculus, with some promising results using simulated students. Rowe and Lester (2015) focused on step-loop adaptivity in the context of an educational game for science learning (CRYSTAL ISLAND; Meluso, Zheng, Spires, & Lester, 2012). They used reinforcement learning to learn a policy by which the system could select the next step in the game's unfolding narrative, in a manner responsive to the state of the narrative and problem solving, as well as to students' prior knowledge and how frequently they play video games. Their study shows that compared to randomly selecting tutorial decisions, the machine-learned adaptive policy had a positive influence on student behavior in the game (greater efficiency of hypothesis testing and

information gathering), though not on domain content learning. A related study (Lee, Rowe, Mott, & Lester 2014) found an improvement in domain content learning over a minimal guidance control condition. None of these projects appear to have shown enhanced learning over a stringent and challenging control condition, due to machine-learned step-loop adaptivity, although that may just be a matter of time.

Discussion of Adapting to Students' Strategies and Errors

We found a substantial number of effective design loop adaptations to students' strategies and errors, more than we expected, underlining once more that use of data from a system can be of significant help in improving that system. We see many demonstration of effective step-loop adaptations to students' strategies and errors. In particular, the use of step-level feedback has been shown to be very effective; for next-step hints, the evidence is more equivocal but on balance, positive (Aleven et al., 2016). Efforts to use machine learning to adaptively select tutoring moves in the step loop have yielded promising results, although not yet a clear demonstration that they enhance learning over a strong, non-adaptive way of selecting these moves. Although we found examples of systems' adapting to student errors or strategies in their task loop, somewhat to our surprise, we did not find any studies rigorously demonstrating the value of such adaptivity, in contrast to the adapting to students' knowledge growth, discussed in the previous section. Perhaps it makes sense that adaptations to faster phenomena (strategies and errors) might need to be faster (i.e., in the step loop) than adaptations to slightly slower phenomena (knowledge growth). Nevertheless, it is plausible that both step-loop adaptations to knowledge growth and task-loop adaptations to strategies and errors could be effective, so we look forward to further research in this area.

Adapting to Affect and Motivation

In this section we discuss empirical evidence regarding whether learning technologies might be more effective if they adapt to, or be adapted to, aspects of student affect or motivation. Work in the area of affect-aware learning technologies (D'Mello, Blanchard, Baker, Ocumpaugh, & Brawner, 2014; D'Mello & Graesser, 2014) tends to focus on affective states such as boredom, confusion, frustration, engagement/flow/engaged concentration (often called "non-basic emotions") as they have been found to be frequent during learning with technologies (e.g., D'Mello, 2013). A number of theoretical frameworks clarify the role of emotions arise in academic settings (e.g., Boekaerts 2007; Pekrun, Frenzel, Goetz, & Perry, 2007), but tend to provide little guidance in designing adaptive learning technologies Nonetheless, we see interesting work in this area.

Adapting to individual differences in affect requires being able to detect a learner's affective states, preferably in a temporally fine-grained and unobtrusive manner. Much work has focused on how machine learning or statistical methods can be used for this purpose. Some methods rely only on data from the regular student-system interaction stream (e.g., Baker et al., 2012) whereas others require data collected with special sensors to measure facial expressions, variables extracted from speech, EEG, heart rate, skin conductivity, pupil dilation, posture, and so forth (Arroyo et al., 2009; Calvo & D'Mello, 2010; Conati, 2002; Conati & MacLaren, 2006; Harley, Bouchet, Hussain, Azevedo, & Calvo, 2015; Sabourin, Mott, & Lester, 2011). The details of this work are beyond the scope of this chapter.

In addition to work on affect-aware learning technologies, a few studies focused on how to adapt to student motivation. There is great variety in theoretical perspectives on the role of motivation in education (e.g., Schunk, Pintrich, & Meece, 2008). However, as in the case of

affect, existing theories are often not specific enough to provide much guidance for technology design. Nonetheless, we found some studies that demonstrate effective ways of adapting to individual students' specific motivations.

Design-Loop Adaptations To Affect and Motivation

A typical design-loop adaptation to affect or motivation might involve the following steps: First, a learning scientist or learning engineer collects data about student affect while students are using a given tutoring system or online course, for example, through systematic observation or frequent polling in the software. Second, she uses the affect data, combined with system log data, to identify parts of the instruction that are particularly frustrating or confusing for large numbers of students (i.e., that tend to induce negative affect or affective states whose frequency correlates negatively with learning gains). She then redesigns these parts of the system in an attempt to make them less frustrating or confusing. Finally, she tests if the new system is more effective than the original version. As a slight variation on this process, she might check which parts of the system are associated with greater incidence of positive affect (e.g., engaged concentration) or greater incidence of affective states that correlate positively with learning. She might then test whether using these features in other parts of the system makes the system as a whole more effective. A small number of studies illustrate this pattern.

Specifically, a series of studies on designing for confusion (D'Mello, Lehman, Pekrun, & Graesser, 2014; Lehman et al., 2013) provide an interesting illustration of design-loop adaptivity to student affect. This work was grounded in prior data analysis that found that confusion can correlate positively with learning (D'Mello et al., 2014), consistent with theories focused on cognitive disequilibrium as being instrumental in learning (Piaget, 1952). These researchers asked, could designing instructional conditions that induce (and resolve) confusion be a

successful instructional design strategy? They created a system capable of conducting *trialogs* (with a natural language speech interface) in which a human student, computer tutor agent, and computer student agent reason through a challenging question in the targeted task domain. (e.g., research design). The trialogs were designed to cause confusion on the part of the human learner by having the two computer agents contradict each other or express false information, later resolved during the trialog. In two studies, confusion was measured as hesitations and pauses in decisions, through self-report, and facial expressions. These studies compared four experimental conditions, which differed only in whether each agent expressed incorrect information. (Thus, there were two conditions with contradictions, and two without.) It was found that contradiction led to deeper comprehension but only if the learners were confused. The work illustrates that adapting to affect in the design loop can be effective. We view this work as an instance of design-loop adaptivity, because the design of the trialogs was grounded in analysis of affect data from prior studies and because the system redesign may have been based on an analysis of dialogue data from a prior system version (e.g., regarding issues about research design may have been confusing to students).

Work by Baker et al. (2009) follows the first part of the prototypical design loop scenario outlined above, focused on a phenomenon he dubbed *gaming the system*, a set of ostensibly disengaged student behaviors. Specifically, Baker et al. (2013) define gaming the system as behaviors in which learners try to get through problems with minimal effort, taking advantage of software features such as hints and step-level feedback to coax answers out of the system without much cognitive effort. This behavior is associated with lower learning when it occurs on problem steps for which the student has a low level of knowledge (Baker et al., 2013). Some work has focused on preventing or counteracting gaming. Using statistical techniques, Baker et al. (2009)

created a classifier capable of detecting gaming behavior automatically based on the regular

student-system interaction stream. This detector was used to investigate which features

commonly found in tutors (especially in the user interface) might make them prone to gaming.

The researchers created an extensive taxonomy of tutor features and used it to code the features

of a large number of tutor units for which they had log data. Running the gaming detector over

the log data from these tutor units enabled them to identify features associated with greater

incidence of gaming. The study provides a foundation for further investigations into design-loop

adaptivity. For example, one might redesign one or more tutor units so they avoid gaming-prone

features identified in this work, and then run a classroom study to test the hypothesis that the

redesigned tutor leads to less gaming and other desirable outcomes (e.g., greater engagement and

learning).

Task-Loop Adaptations To Affect and Motivation

Some advanced learning technologies can adapt to affect and motivation in their task loop, with a

positive influence on student learning. One project focused on adaptations to elicit students'

personal interest (Hidi & Renninger, 2006). Past research shows that higher interest is associated

with learning (Ainley, Hidi, & Bendorff, 2002; Harackiewicz, Durik, Barron,Linnenbrink-

Garcia, & Tauer, 2008; Hulleman & Harackwicz, 2009). One way to elicit interest is to

personalize instructional contexts according to students' out-of-school interests (e.g., sports,

gaming, movies; Anand & Ross, 1987; Cordova & Lepper, 1996). Walkington and colleagues

investigated the effect of adjusting the cover stories of algebra problems to students' personal

interest, in a study with 145 students in a Cognitive Tutor Algebra course (Walkington, 2013;

Walkington & Sherman, 2012). Students randomly assigned to conditions received either normal

problems or personalized problems based on their personal interest in sports, music, art, or

games. The mathematics in the assigned problems was the same in both conditions. The results were quite remarkable. Personalization based on interest improved the immediate accuracy and efficiency of learning and led to accelerated future learning with the tutoring software, four units later, without the personalization. The effects were strongest when normal problems were disconnected from student experiences (e.g., nitrogen in an asteroid) rather than already somewhat personalized (e.g., money at work). The effects were largest for students identified as struggling with Algebra I. In a follow-up study with 152 students, a similar personalization intervention led to increased triggered situational interest when working in the software, improved accuracy and efficiency in the tutor, and enhanced individual interest in learning mathematics, which was associated with gains on paper-based tests of Algebra I skills (Bernacki & Walkington, 2014; Walkington & Bernacki, 2015).

Similarly, a study by Heilman et al. (2010) found that learners of English as a second language learn better when reading materials are selected (by machine algorithms) in a manner that takes into account not just cognitive, domain-based factors (e.g., what vocabulary words the learner has not learned yet, and the difficulty level of the text) but also based on the learner's personal interest.

Finally, work by Baker et al. (2008, 2013) provides some evidence for the effectiveness of adapting, in a system's task loop, to students' gaming the system. As mentioned, Baker et al. created a machine-learned detector that automatically detected the occurrence of disengaged gaming behaviors. They embedded the detector in a tutoring system for middle-school mathematics, so that, when it detected a student's gaming of the system, it would assign learning tasks that focused on prerequisite knowledge for the instructional objectives targeted in the given unit. For students prone to gaming, this task-loop adaptation reduced gaming behaviors and led

to improved learning, compared to working with the regular tutor version, in which there was no adaptivity to student gaming.

Step-Loop Adaptations Affect and Motivation

A number of studies have investigated the effectiveness of step-loop adaptations to affect. These studies tended to use conversational agents and other kinds of tutorial dialogue systems. These systems interact with students in natural language (sometimes in speech, sometimes through a chat interface) and often show an animated agent on the screen that represents a tutor or peer learner character.

One study evaluated the effect of empathic dialogue moves by a pedagogical agent (AutoTutor). The affect-aware AutoTutor system selected the moves adaptively based on the student's affective state (D'Mello et al., 2010). The system was capable of detecting boredom, confusion, frustration, and neutral affective state using discourse features, body language (posture), and facial expressions, for which it required some special sensors not part of a standard computer configuration. It responded with empathetic, encouraging, motivational dialogue moves and emotional displays. For example, the tutor agent might say: "This stuff can be kind of dull sometimes, so I'm gonna try and help you get through it. Let's go!" or: "Some of this material can be confusing. Just keep going and I am sure you will get it." In a study with 84 participants, students were randomly assigned to an affect-aware version of AutoTutor or the regular version, whose dialogue moves were selected based on the state of dialogue, but without responding to student affect. The affect-aware empathic dialogue moves enhanced learning for lower domain-knowledge students, who also perceived the affect-aware system more positively. We consider the empathic dialogue moves to be step loop adaptations, because they adjust what happens on the specific steps within a problem (i.e., turns within a dialogue).

A study with a system called Gazetutor (a pedagogical agent that gives lectures) also demonstrated that adaptively responding to student affect in a system's step loop can be efffective (D'Mello, Olney, Williams, & Hayes, 2012). During the lecture presented by the pedagogical agent, this system used eye tracking to detect disengagement on the part of the student as not looking at relevant areas of the screen for 5 seconds or more. It addressed students in disengaged state using prompts such as "Please pay attention" or "You might want to focus on me for a change." The gaze-sensitive tutor helped students re-orient attention and learn with greater understanding, compared to a version that did not respond to disengagement detected through eye tracking.

Finally, Forbes Riley and Litman and (2011) tested whether, in a tutorial system that engages in natural language dialogue with the student, adapting to student uncertainty as detected in the speech signal leads to better learning. The system treated correct but uncertain responses differently from correct but certain responses, in a manner similar to the study by Mitrovic et al. (2013), described above. For correct responses about which the student appears to be certain, the dialogue agent merely acknowledged that the answer is correct. When the answer was correct but the student was uncertain, however, the system acknowledged that the answer was correct and then elaborated on why, in the same manner as it did for incorrect responses, either by providing an explanation or by leading the student through a line of reasoning. The results of an evaluation study were inconclusive, however, possibly due to the relatively low incidence of correct but uncertain responses (which were the only occasions for adapting to uncertainty).

Discussion of Adapting to Student Affect and Motivation

The work reviewed here shows that adapting to student's affect can be effective in all three loops (design loop, task loop, and step loop. This line of work is instrumental in helping us

understand how affect interacts with academic learning. There is a much larger space of affect-sensitive adaptations to be explored. An interesting question is how much guidance studies of human tutors' responses to affect might provide (Lehman, Matthews, D'Mello, & Person, 2008). Also, it appears that female and male students react differently to affective adaptive support (Vail, Boyer, Wiebe, & Lester, 2015).

Although less work has focused on adapting to student motivation than on adapting to student affect, the limited work that has been done in this area shows that selecting tasks based on personal interest can have a strong effect on student learning. Some work (not discussed above) has focused on automatically detecting motivational states (specifically, self-efficacy) using machine-learned models (McQuiggan, Mott, & Lester, 2008). These methods could be used for adaptive technologies that respond adaptively to students' changing self-efficacy.

Some work emphasizes the interplay between affect and cognitive factors. For example, Baker et al.'s (2008) tutor reacted with a cognitive intervention (namely, remedial practice) to disengaged behaviors. Along similar lines, some researchers are now designing adaptive algorithms for task selection that consider affect in combination with cognitive factors such as knowledge growth or problem-solving success. For example, Arroyo et al. (2014) describe an effort-based adaptive tutoring algorithm that integrates engagement factors, including a measure of the effort the student is exerting in the learning process. Mazziotti et al. (2015) present an algorithm for task selection that integrates affect and cognitive factors. In short, we are seeing forms of hybrid adaptivity that cut across the rows of the Adaptivity Grid (i.e., that integrate multiple psychological realms).

Adapting to Students' Self-regulation of Learning

Self-regulated learning (SRL) refers to self-directive processes and associated motivational beliefs that enable learners to take a proactive role in acquiring academic skill, by setting goals, selecting and deploying strategies, and self-monitoring one's effectiveness (Zimmerman, 2008). Theories of SRL abound (e.g., Winne & Hadwin, 1998; Zimmerman & Campillo, 2003; Zimmerman, 2008)' SRL accounts for substantial variability in learning outcomes (Zimmerman & Martinez-Pons, 1988, 1990;,Pintrich & De Groot, 1990). Here we consider the question of whether advanced learning technologies might be more effective if they adapt to learners' self-regulation of learning. Adapting to self-regulation means that data is used to design support that helps students in applying self-regulatory processes during their learning activities. Sometimes, the main goal is to help students learn more effectively at the domain level, helped by the support for SRL. Sometimes, a key additional goal is to help students become better at regulating their own learning so that their *future* learning experiences, when the support for SRL is no longer in effect, will be more effective. We discuss projects focused on both types of goals.

Design-Loop Adaptations To Self-Regulated Learning

A typical design-loop adaptation to students' self-regulation of learning would involve analyzing data from a given system to better understand student self-regulation with that type of system and/or to find evidence of productive or not-so-productive forms of self-regulation. This data mining and modeling phase would be followed by a redesign of the system to better support the identified self-regulation processes, and/or the learning thereof. The redesigned system could be adaptive either to student similarities with respect to the targeted SRL processes.

A considerable amount of work has focused on supporting self-explanation within adaptive learning technologies (Corbett, Wagner, & Raspat, 2003; McNamara, O'Reilly, Rowe,

Boonthum, & Levinstein, 2007; Rau, Aleven, & Rummel, 2009; Weerasinghe & Mitrovic, 2006; Wylie, Sheng, Koedinger, & Mitamura, 2011). Some of this work can be viewed as design-loop adaptivity. For example, a set of studies on supporting self-explanation with a Cognitive Tutor (Aleven & Koedinger, 2002) started with a discovery in data, a key criterion for design-loop adaptivity. On pre- and post-tests, students were better able to solve geometry problems than provide reasons for their problem steps in terms of geometry theorems and definitions (Aleven, Koedinger, Sinclair, & Snyder, 1998). This finding suggested that students learned shallow strategies, such as relying on vague notions of symmetry in geometry diagrams or relying on the fact that angles look the same in the diagram. It led to the hypothesis that prompting students to articulate the reasons behind their steps during their work with the tutor might lead to more robust learning, in line with work on self-explanation at the time (e.g., Chi, Bassok, Lewis, Reimann, & Glaser, 1989; Chi, de Leeuw, Chiu, & LaVancher, 1994; Renkl, Stark, Gruber, & Mandl, 1998). The tutor was redesigned so that it prompted students to explain problem steps by providing the name of the geometry theorem that justified the step. Two classroom studies compared students' learning outcomes with the redesigned tutor versus the original tutor. The original tutor provided guidance during problem solving but did not offer any prompts or support for self-explanations. The redesigned tutor was the same, except that prompted for explanations and provided hints and feedback regarding these explanations. The support for self-explanation led to stronger conceptual knowledge and less shallow procedural knowledge (Aleven & Koedinger, 2002). We consider this work as an instance design-loop adaptivity, because the support for the targeted SRL process was designed based on data from the given system and the support was the same for all learners.

Task-Loop Adaptations To Self-Regulated Learning

A typical task-loop adaptation to student SRL would be a system that supports student self-regulation in-between tasks (e.g., provides support for making good task selection decisions, known to be challenging for students) or that tries to select tasks that pose specific challenges regarding SRL. Although the latter approach is interesting, we only know of work that has taken the former approach. This work has focused on the use of "open learner models" to adapt to or support student SRL. Open learner models are a common feature in intelligent tutoring systems. They are presentations of a system's assessment of student characteristics such as their current level of knowledge related to instructional objectives, often shown in a convenient graphical format (Bull & Kay, 2010; Corbett & Anderson, 1995; Mabbott & Bull, 2004). Open learner models often take advantage of the system's underlying student modeling technologies (e.g., Desmarais & Baker, 2012). They can serve a variety of purposes; one of their functions has long been considered to be "promoting metacognitive activities such as reflection, planning and self-monitoring" (Bull & Kay, 2010, p. 301), typical self-regulatory processes. A small number of empirical studies have tested this proposition.

Arroyo et al. (2007, 2014) found benefits of periodically presenting a simple student model in the task loop of an intelligent tutoring system. Specifically, after every six problems, their system (Wayang Outpost, for middle-school mathematics) presented progress charts to the student, which documented his/her own recent domain-level learning and performance, accompanied by brief metacognitive tips that encourage good study habits. A pseudo-experimental classroom study with 88 students compared learning with a tutor version with progress charts and tips to learning with a tutor version that did not have progress charts or tips, but that was otherwise the same. The progress charts and tips influenced student behavior with fewer quick guesses at answers and greater engagement, as indicated by spending more time per

problem on subsequent problems. Also, they led to higher pre/post learning gains and, remarkably, a higher passing rate on state exams.

Studies by Long and Aleven (2013a; 2013b) investigated how and how well open learner models can support self-assessment, a key self-regulatory process. In one classroom study, with an intelligent tutoring system for high-school geometry, these researchers tested benefits of periodically filling out paper skill diaries with self-assessment questions related to the learner's own open learner model, compared to filling out a "control diary," which asked superficial questions that did not involve self-assessment. Especially lower-performing students benefitted from the skill diaries; they had better learning outcomes than their counterparts in the control group. Also, the accuracy of their self-assessment improved from pre-test to post-test. A subsequent study, this time with Lynnette, an intelligent tutoring system for middle-school equation solving, found that an open learner model with added prompts for self-assessment (this time in the software, rather than in paper diaries) can help students attain better learning outcomes (Long & Aleven, 2013b), compared to working with a tutor version without an open learner model or self-assessment prompts. In a subsequent, larger, study, this result was replicated although it was restricted to students who had some amount of control over problem selection (Long, 2015).

Finally, a series of studies by Mitrovic and Martin (2007) found mixed evidence regarding whether an open learner model can have a positive effect on students' task selection decisions. Generally, these studies suggest that adapting to student self-regulation in a system's task loop can be effective, especially when an open learner model in combination with additional support for self-assessment and learner-controlled task selection.

Step-Loop Adaptations To Self-Regulated Learning

A number of projects provide evidence that tutor agents that respond to aspects of student self-regulation within a task (i.e., within the system's step loop) can be effective. An early approach was called intelligent novice tutoring (Mathan & Koedinger, 2005). The purpose was to help students learn a self-regulatory strategy for recognizing and fixing errors, aided by "grounded feedback" (Nathan, 1998; Stampfer & Koedinger, 2013). Grounded feedback is feedback that reflects back consequences of actions in a natural or easily-understood way, rather than indicating only whether the action is correct or not. The given tutor helped students learn to use and copy formulas in Excel; the output of Excel formulas, shown by Excel served as grounded feedback (e.g., the numbers resulting from applying a formula may be outside of the expected range, or error codes may be displayed). When a student made an error, the intelligent novice tutor intervened only when the student moved on to other things without correcting the error – that is, when the student made an error with the targeted self-regulatory strategy of error detection and correction. It then helped the student with error diagnosis and correction. This approach improved students' domain-level learning, compared to immediate correctness feedback without the grounded feedback (Mathan & Koedinger, 2005). Thus, the experiment demonstrated that step-loop adaptation to SRL can enhance domain-level learning.

Since then, a number of tutoring approaches to SRL have been developed and evaluated. Roll and colleagues developed a tutor agent that provides adaptive feedback on students' help-seeking behavior with an intelligent tutoring system (Aleven, McLaren, Roll, & Koedinger, 2010; Roll, Aleven, McLaren, & Koedinger, 2011). This tutor agent was able to recognize, in a context-sensitive manner, many help-seeking errors, such as using hints to get answers without understanding, or not requesting a hint when a hint would objectively appear to be highly useful (e.g., after multiple errors on a step) (Aleven & Koedinger, 2000; Aleven, McLaren, Roll, &

Koedinger, 2006; Aleven, Roll, McLaren, & Koedinger, 2016). Two classroom studies compared

learning with two versions of a tutoring system for geometry learning that were the same except

one gave feedback on help seeking, in the context of geometry problem solving. The second

study found a lasting improvement of students' help-seeking behavior, due to feedback on help

seeking, although no effect on domain-level learning (e.g., Roll et al., 2011) was found, due to

improved help seeking.

A related project by Azevedo and colleagues, called Meta-Tutor, provides evidence for

the effectiveness of automated tutoring on a broader range of SRL processes (Azevedo et al.,

2012). This project created tutor agents that support key SRL processes in a hypermedia learning

environment through prompts and feedback. These processes included setting sub goals, writing

a summary, assessing how relevant particular content is, taking notes, assessing one's own

understanding, and so forth. In an experiment with 83 students, students who worked with a

version of the hypermedia system that gave tutoring support for SRL achieved better learning

outcomes than students who worked with the hypermedia system without this support, although

the effect was found only for students with a performance approach orientation (Duffy &

Azevedo, 2015).

Finally, studies with a learning environment called Betty's Brain tested the value of

adaptive tutoring on aspects of SRL in the context of learning by teaching (Leelawong & Biswas,

2008). The student's task was to teach Betty, a teachable agent, by building a concept map that

captures causal processes in a science domain. To help build the map, students could read

available textual sources. Key self-regulatory strategies in this environment are information

seeking and solution quality monitoring by requesting and interpreting feedback. Studies with

5th-grade students compared learning in the standard teachable agent environment against

learning with a version of that environment with adaptive support for the targeted SRL

processes. In the standard teachable agent environment, feedback focused on the correctness of

the student's concept map and included explanations of errors in this map. By contrast, in the

SRL version of the environment, students were given feedback and advice regarding information

seeking and monitoring map quality. A second study found that this support helped students do

better on a "preparation for future learning" (PFL) task, in which the students created a new

concept map with minimal support (Tan, Biswas, & Schwartz, 2006) These students learned

better information seeking and monitoring skill that enabled them to make better concept maps

(Wagster, Tan, Wu, Biswas, & Schwartz, 2007). No pre/post test of domain content was given,

however. A similar result was obtained in a later study (Kinnebrew, Szegedy, & Biswas, 2014).

This work was therefore one of very few projects, together with Roll et al. (2014), that showed

that effects of support for SRL can persist even when it was no longer in effect.

Discussion of Adapting to Student Self-regulated Learning

We find successful approaches to adapting to SRL in all three loops, although we also

find some inconclusive results. A number of studies provide evidence that design-loop

adaptations to SRL can be effective. These studies focused on supporting self-explanation. A

number of studies tested whether an open learner model in an intelligent tutoring system can

support effective task selection and self-assessment, key self-regulatory processes. The evidence

is somewhat mixed; more work remains to be done to understand the circumstances under which

an open learner model can support effective self-regulation, how best to design this kind of

support, and whether it can have a lasting and transferable effect (e.g., Roll et al., 2014). A

number of projects demonstrate that step-level tutoring of SRL can help improve students' SRL

processes, including demonstrations that the improvement can last beyond when the support is in

effect (Roll, Aleven, McLaren, & Koedinger, 2011; Tai, Arroyo, & Woolf, 2013; Tan, Biswas, & Schwartz, 2006;). Similarly, some projects found that domain-level learning outcomes improve due to adaptive support for SRL, illustrating the promise of adaptively supporting SRL in learning technologies (Aleven & Koedinger, 2002; Mathan & Koedinger, 2005; Duffy & Azevedo, 2015). However, a number of projects failed to find an effect on domain-level learning. Few projects have tested effects on future learning. Improving future learning (at the SRL level and at the domain level) remains as an important challenge. While there have been some interesting successes, work remains to be done to demonstrate the generality and transferability of the SRL models underlying adaptive learning technologies. Also, it will be interesting to link explicitly to theoretical frameworks of SRL (e.g., Aleven, 2013).

Adapting to Student Learning Styles

Finally, should instruction adapt to students' learning styles? There is a very large literature as well as substantial controversy regarding this topic, as it pertains to both computer-based and other learning environments. We follow Pashler, McDaniel, Rohrer, and Bjork (2008), who define learning styles as "the concept that individuals differ in regard to what mode of instruction or study is most effective for them." We note that others (e.g., Jonassen & Grabowski, 1993, p. 233) define learning styles more narrowly as *self-reported preferences* for certain ways of processing information. Pashler et al. do go on to say, however, that [italics added] "assessments of learning style typically ask people to evaluate what sort of information presentation they *prefer*" and comment further that the notions of learning styles as preferences versus abilities are often conflated. One reason that research in this area is challenging is that there are many and varied taxonomies of learning styles—a review in 2004 covers 71 of them (Coffield, Moseley, Hall, Ecclestone, & Vorhaus, 2004). We also note that researchers

commonly construe learning styles as a relatively stable trait. Therefore, it is typically measured only once, at the outset of the instruction. This single measure is then used to adapt subsequent instruction accordingly. In contrast, we are increasingly seeing technology-based efforts to assess learning styles dynamically, during the learning process, by creating automated methods that infer students' learning styles from behavior displayed in interactions with computer-based learning environments.

The notion that instruction is more effective when it aligns with each student's learning style has substantial intuitive appeal (Howard-Jones, 2014; Willingham, Hughes, & Dobolyi, 2015). Howard-Jones (2014) for example, reports that a survey of over 932 teachers from five different countries revealed that over 96% of teachers (with the percentage in each country above 90%) believed that students learn better when information is presented in a way that matches their learning style. There is however substantial controversy about the questions of whether, when, and how adapting to students' learning styles can improve student learning. A number of prior reviews, notably one by Pashler et al. (2008), concluded that there was scant evidence to support the notion that adapting to learning styles makes instruction more effective. They did not find a single study that demonstrates a crossover interaction between learning styles and methods of instruction in terms of learning outcomes. Several studies cited there do use such methodology but fail to find any interaction (Constantinidou & Baker, 2002; Cook, Thompson, Thomas, & Thomas, 2009; Massa & Mayer, 2006). This standard of evidence and the dearth of studies meeting this standard are restated more recently in Rohrer and Pashler (2012). A later review by Kirschner and van Merriënboer (2013) came to the same conclusion, as does Howard-Jones (2014) and Willingham, Hughes, and Dobolyi (2015). These reviews support the conclusion that adaptation to learning styles is largely unproven and lacks a theoretical basis.

These reviews however did not look at the large body of literature regarding the influence of learning styles in the context of learning technologies, primarily adaptive educational hypermedia systems (AEHs) (Brusilovsky, 2001). This term denotes online instruction that offers multiple media (e.g., videos, text) and, typically, freedom of navigating the learning materials. In this large literature, a slightly different, though still conflicted, picture emerges. A review specific to computer-assisted instruction, with a (non-exclusive) focus on medical domains, concluded that "if aptitude–treatment interactions with CLSs [cognitive/learning styles] exist, they seem to be infrequent and small in magnitude" (Cook, 2012). Of 65 contrasts in the reviewed studies, only 9 (14%) yielded a statistically significant aptitude-treatment interaction. This review therefore largely confirms the conclusions by Pashler et al., but would seem to leave the door open for learning style adaptivity to be occasionally useful, since a small percentage of reviewed studies showed an interaction between instructional treatment and learning style, with respect to students' learning outcomes. Nonetheless, Cook, after unsuccessfully trying, in his own research, to confirm the notion that adaptivity to learning styles is helpful, concludes that the effect of learning styles, if it exists at all, is often overwhelmed by the effect of the instructional method.

Two recent reviews came to somewhat more positive conclusions regarding the value of adapting to learning styles in learning technologies (Akbulut & Cardak, 2012; Özyurt & Özyurt, 2015), though, in our opinion, without a careful weighing of the evidence for and against. Akbulut and Cardak (2012) reviewed 70 articles including both technical and empirical papers. Over half the empirical papers were "design-based case studies," which do not answer our central question, whether adapting to learning styles makes instruction more effective. Regarding empirical studies, they concluded: "When concrete learning outcomes in robust experimental

studies were considered, findings were slightly controversial. That is, there were eight studies indicating that AEH systems significantly affected the learning outcomes of the students in a positive way, and there were four studies refuting the significant effects on the learning outcomes." Özyurt and Özyurt's (2015) review of literature concerning adapting e-learning to learning style included 69 studies published between 2005 and 2015, of which 22 were concerned with the effect on academic achievement. They concluded that "[t]hough positive effects of AEHs on academic achievement and learning outputs were not clearly revealed in some experimental studies, majority [sic] of the studies yielded positive results in this matter." . It is unfortunate however that neither of the Akbulut and Cardak (2012) and Özyurt and Özyurt (2015) reviews attempt to evaluate the methodological rigor of the studies that were considered, reconcile the conflicting results in the literature, or narrow down the range of circumstances under which adapting to learning styles may be helpful. Thus, we ask whether the work cited in these reviews force us to reconsider the conclusions reached in the earlier reviews by Pashler et al. (2008) and by Kirschner and Van Merriënboer (2012). Our answer is that only a small number of rigorous studies show advantages of adapting to learning styles. We review some of them below. We focus on task-loop adaptations to learning styles as we did not find any work on adaptations to learning styles in the design or step loop.

Task Loop Adaptation to Learning Styles

Ford and Chen (2001) studied the value of adapting instruction to the students' level of "field dependence."   Field dependence/independence denotes "the extent to which the organization of the prevailing field dominates perception of any of its parts" (Witkin, Oltman, Raskin, & Karp, 1971, p. 7). This learning style is typically measured through a performance test rather than self-report. According to prior work, field independent learners are better at analytic

activity and imposing structure on (relatively) unstructured information; they favor a hypothesis-testing approach to learning. Field dependent learners, by contrast, are not as good at imposing structure and favor a more pre-structured learning environment. Ford and Chen hypothesized that field-independent learners might do better in a learning environment in which the topics are sequenced in a breadth-first manner, which tends to support a "big picture before details" approach, whereas field-dependent learners might do better in a learning environment in which topics are sequenced in a depth-first manner, which conforms to a "topic-by-topic" approach. They tested this interaction hypothesis with 73 postgraduate students in the context of an online tutorial on HTML, studied for 1.5 hours. Based on their learning styles, students were assigned either to a "matched" or a "mismatched" condition, with materials sequenced in either a breadth-first or depth-first manner. The hypothesized interaction was confirmed: Field-independent learners showed greater learning gains with the breadth-first version of the materials, whereas field-dependent learners showed greater learning gains with the depth-first version of the materials. This interaction occurred only for male participants (there were roughly an equal number of male and female participants) and only with respect to conceptual knowledge of HTML, as assessed by multiple-choice questions, not in an application task. In short, the study finds better learning outcomes, due to adapting to learning styles, for a subset of students on a subset of measures.

A study by Popescu (2009; 2010) like the Ford and Chen study, studied effects of changing the recommended sequence of materials, though based on a variety of learning style classifications. This study found that a matched condition (in which the sequence of the materials matched the students' learning style) leads to slightly greater efficiency and enjoyment, compared to a mismatched condition, though not to greater learning.

Several studies compared instruction that adapts to learning styles (i.e., a matched condition) not to a mismatched condition, but to a non-adaptive condition. A study by Mampadi, Chen, Ghinea, and Chen (2011) tested the value of adapting to a Holist versus a Serialist learning style (Pask, 1976). Holists prefer to gain a view of the big picture first (e.g., important concepts and their connections), with details to be filled in gradually, whereas Serialists tend to prefer a topic-by-topic approach. The study involved an adaptive educational hypermedia system for learning about XML, a widely used file format in computer science. Students' preferences for a Holist or Serialist style were assessed using the Study Preference Questionnaire (SPQ), a five-item self-report questionnaire (Ford, 1985). Students in the experimental condition who preferred a Serialist style of learning were assigned to the Serialst version of the system whereas students in the experimental condition who preferred a Holist style of learning were assigned to the Holist version of the system. The Serialist system version was designed to support sequential access to the learning materials offered on the site. It offered next/previous buttons to navigate the materials, an alphabetical index of the site's content, and no support for jumping around (i.e., it restricted navigation choices and offered no links in body text). By contrast, the Holist version of the system offered no direct guidance for a particular sequence through the materials. Rather, it was designed to let students jump around, with many links in body text, and a hierarchical map of the site content. The students in the control condition used a system version that combined the features of both the Serialist and the Holist version of the system. Learning was assessed with a pre- and post-test with multiple-choice questions regarding XML. It was found that among students with a preference for a Holist learning style, those that used the system version geared toward a Holist learning style had higher post-test scores and higher gain scores (post-test minus pre-test) than those who used the control version (i.e., the full-featured version). The analogous

result was found for the Serialists. Further, students in the adaptive condition perceived their respective system version more positively than did their counterparts in the control condition. Thus, catering to a student's learning style was found to be better than a non-adaptive approach that offers all options to all students. We have some reservations about this study and its write-up, however. The paper does not state how participants were recruited or assigned to conditions, so we do not know if random assignment was used. Also, because the control is not the same instruction but without adaptation, we cannot isolate the cause of the observed benefit as adaptivity. Finally, it is hard to know whether the system version used in the control condition represents a bona fide or ecologically valid AEH, or whether it perhaps imposes high cognitive load by having too many features.

A study by Graf and Kinshuk (2007) with 235 university students investigated the value of adapting an online course (on object-oriented modeling, a basic computer science topic) based on three dimensions in the Felder-Silverman learning style model (FSLSM) (Felder & Silverman 1988), namely, the active/reflective, sensing/intuitive, and sequential/global dimensions. For each of these dimensions, the online course varied the sequence and number of course elements of various types, such as chapter outlines, chapter conclusions, regular content pages, examples, self-assessment tests, and exercises. For example, active learners received more exercises and self-assessment tests, but fewer outlines and conclusions/summaries, compared to reflective learners. As another example, for sensing learners, examples were presented after the related abstract materials, whereas the order was reversed for intuitive learners. The dimensions were assessed at the beginning of the instruction using the ILS questionnaire, a 44-item questionnaire developed by Felder and Soloman (1997). Students were assigned randomly to three conditions working with respectively a course version that matched their learning style (matched),

mismatched their learning style (mismatched), or contained all materials (standard). The study

found more efficient learning (with no difference in learning outcomes) for students in the course

version that matched their learning style, compared to students in the mismatched version or the

standard version. Unfortunately, the paper does not report the extent of the efficiency difference,

only that it was statistically significant.

A similar effect of adapting to learning styles was found in a study by Tseng, Chu,

Hwang, and Tsai (2008). Although there were no differences in students' learning outcomes,

students in the version that adapted to learning styles spent less time, compared to those working

with a system version that did not adapt. Other studies however failed to find a difference in

adapting to learning styles (e.g., Brown, Brailsford, Fischer, & Moore, 2009).

Finally, some studies that are sometimes cited as supporting the value of adapting to

students' learning styles tested an adaptive method that took into account not only students'

learning styles, but also an assessment of student knowledge (Despotovic-Zrakic, Markovic,

Bogdanovic, Barac, & Krco, 2012; Limongelli, Sciarrone, Temperini, & Vaste, 2009; Mustafa

& Sharif, 2011), a point made also by Popescu (2010). From the point of view of creating

effective adaptive learning technologies, it makes perfect sense to adapt to student knowledge in

combination with other student characteristics. However, these types of experiments do not help

in testing the value of adaptivity specifically to learning styles.

Discussion of Adaptation to Learning Styles

Although we discussed some studies that found beneficial effects for adapting to learning styles,

in the end, we do not diverge far from the conclusions by Pashler et al. (2008), Kirschner and

van Merriënboer (2013), and Cook (2012), that adapting to learning styles is not often effective.

Cooks's (2012) viewpoint, that the instructional method often matters more than learning style,

resonates with us. Further, we are more drawn to look at the question of how learners can be helped in operating effectively in given learning environments through the theoretical lens of self-regulated learning. This entails identifying what learning strategies work in a given kind of environment, using the tools of cognitive task analysis, and studying what instruction can help learners develop these strategies. While this approach has its own challenges, it is grounded in what works, rather than in student preferences, and might help get us away from a view of learning styles as traits, which would be more in line with view of strategic flexibility as an important element of expertise, and in line with views of intelligence as malleable.

Conclusion

This chapter gives an overview of the state of scientific knowledge of when and how adaptive instruction can be effective in enhancing educational outcomes. Our review focuses on empirical evaluations of educational effectiveness of adaptive technologies as compared to non-adaptive alternatives. We also look at studies that tested for crossover interactions between student variables and instructional treatments. Compared to previous reviews of adaptive instruction, novel features of this review are the emphasis on the use of data as a crucial element of developing adaptive instruction, an Adaptivity Grid that organizes a design space for adaptive instruction, highlighting design-loop changes as a form of adaptivity, and focusing on empirical studies rather than technical aspects. The review focuses on studies of students learning individually. We do not address collaborative learning.

The review spans a large design space, captured in a 3x5 Adaptivity Grid. We distinguish three forms of adaptivity: step-loop adaptivity (within a problem), task-loop adaptivity (between problems), and design-loop adaptivity (between system versions, based on data-driven redesign). We distinguish five broad psychological realms that can be the basis of adaptations: the student's

knowledge, her path through an activity, affect/motivation, self-regulation, and students' learning styles. The grid could be expanded to include additional factors, such as social factors or relatively stable student characteristics such as working memory capacity and spatial reasoning ability.

We found evidence of effectiveness of adaptivity for 13 out of 15 cells of the Adaptivity Grid, albeit of varying strength. Covering this many cells was somewhat of a surprise to us; when we embarked on this review, we did not fully know the range of existing empirical research. (The cells not covered are design-loop and step-loop adaptations to learning styles.) Thus, the reviewed body of work demonstrates a wide range of methods for effective adaptation of instruction. We find some general trends in the series of studies reviewed, which give us insight not only into what works and what can be used now to improve instruction, but also into what might be ahead in terms of further research in this area. The trends are: (a) prevalence of effective design-loop adaptations; (b) strong evidence for adapting to knowledge and knowledge growth in the design loop and task loop; (c) strong evidence for adapting to strategies and errors in both the design loop and step loop; (d) lack of evidence for adapting to learning styles; and (e) emergence of hybrid forms of adaptivity, meaning ways of taking into account multiple psychological realms, such as combinations of knowledge, path through problem, and affect. We discuss each trend in turn.

Prevalence of Effective Design-Loop Adaptations

One surprise to come out of this review is the prevalence and effectiveness of data-driven design loop adaptations. Design-loop adaptations adjust instruction to student similarities, as in the redesign of the statistics course by Lovett et al. (2008). We found effective design-loop adaptations for four of the five psychological realms of the Adaptivity Grid. The prevalence of

design loop adaptations can be viewed as evidence that big data is starting to impact classroom-ready products of educational research. It should also be viewed as evidence that getting instructional design right in a first iteration is more difficult than is often thought, and that instructional design should be approached as an iterative, data-driven process, supported for instance by visualizations and methods for educational data mining. This notion may apply in a broad range of instructional design, not just advanced learning technologies.

Strong Evidence for Adapting to Student Knowledge in the Design Loop and Task Loop

If, as many authors before us have noted, there are big differences in the prior knowledge that students bring to bear, it would follow that instruction that adapts to the knowledge and knowledge growth of individual students will often be more effective than instruction that does not. This notion is borne out by our review, as we found effective demonstrations of adaptivity in all three loops, but particularly the design and task loops. Regarding the task loop, the evidence is strong that mastery learning approaches implemented in technology (such as those used in Cognitive Tutors; Corbett et al., 2000) can be practical and effective in actual classrooms. Perhaps contrary to common intuition, data also shows that oftentimes, the specific difficulties that learners experience in a given task domain are largely the same. This might be one way to explain why design loop adaptations, which cater to similarities among learners, can be effective.

Strong Evidence for Adapting to Strategies and Errors in the Design Loop and Step Loop

The evidence in favor of adapting to strategies and errors in a system's step loop is strong – perhaps no surprise, as the step loop has often been viewed as a strength of intelligent tutoring systems (VanLehn, 2006; 2016) and evidence bears it out (VanLehn, 2011). But it was a surprise that many design-loop adaptations to strategies and errors were shown to be effective. An

interesting result is further that although adaptive technologies have recognized the need for supporting strategy development and strategic flexibility, merely being able to recognize and react to multiple strategies may not be enough (Waalkens et al., 2013).

Lack of Evidence for Adapting to Learning Styles

Evidence that adapting to learners' preferred style of learning helps make instruction more effective is weak, even though there is no lack of work in this area and it is a common belief, even among teachers, that learners learn more effectively when instruction matches their style of learning. Early reviews (e.g., Pashler et al., 2008) argued that very few studies had been conducted that constitute a fully rigorous test of the value of instruction that adapts to learning styles. They also argued that a small number of studies provided evidence against that notion. Since then, additional reviews have appeared that focused on learning technologies. One review came to the conclusion that adapting to learning styles is often (though not always) ineffective, as instructional methods tend to overwhelm the effect of learning styles (Cook, 2012). Two other reviews (Akbulut & Cardak, 2012; Özyurt & Özyurt, 2015) present a more optimistic picture, but without, we feel, a critical appraisal of the evidence across studies or trying to reconcile conflicting results in this literature. Although we found some studies that support the notion that adapting to learning styles can be effective, more studies failed to find an effect. At this point in time, it is very difficult to predict when and how adapting to learning styles is effective.

Prevalence of Hybrid Forms of Adaptivity

Finally, an emerging trend is that the "What to adapt to?" question is increasingly being approached as involving and integrating multiple psychological realms. This focus leads to hybrid adaptive policies that are responsive to a variety of student variables (e.g., Arroyo et al.,

2014; Rowe & Lester, 2015; Mazziotti et al., 2015) across the rows of the Adaptivity Grid. This work is consistent with findings in the literature that the different features to adapt to might interact (e.g., Vail et al., 2015, Moos & Azevedo, 2008). In part, this trend may be inspired by the idea that adaptation should take into account multiple realms, for example, because it is limiting to focus on cognitive factors only. In part, this trend may be inspired by knowledge representation (such as Bayesian Networks) or machine learning approaches (e.g., reinforcement learning) that easily combine information from different realms. At this point in time, we are not aware of any empirical results that show that adapting to multiple psychological realms is better than adapting to a single realm, or to a fixed (i.e., non-adaptive) instructional sequence. However, it is only a matter of time before increasingly complex adaptive systems will emerge that test this and other innovative ideas!

References

Adams, D., McLaren, B. M., Durkin, K., Mayer, R.E., Rittle-Johnson, B., Isotani, S., & Van

    Velsen, M. (2014). Using erroneous examples to improve mathematics learning with a

    web-based tutoring system. *Computers in Human Behavior, 36,* 401-411. doi:

    10.1016/j.chb.2014.03.053.

Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes

    that mediate their relationship. *Journal of Educational Psychology, 94*(3), 545-61.

Akbulut, Y., & Cardak, C. S. (2012). Adaptive educational hypermedia accommodating learning

    styles: A content analysis of publications from 2000 to 2011. *Computers & Education,*

    58(2), 835–842. doi: 10.1016/j.compedu.2011.10.008

Aleven, V. (2013). Help seeking and intelligent tutoring systems: Theoretical perspectives and a

    step towards theoretical integration. In R. Azevedo & V. Aleven (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., Beal, C. R., & Graesser, A. C. (2013). Introduction to the special issue on advanced

    learning technologies. *Journal of Educational Psychology, 105*(4), 929-931.

Aleven, V., & Koedinger, K. R. (2013). Knowledge component approaches to learner modeling.

    In R. Sottilare, A. Graesser, X. Hu, & H. Holden (Eds*.), Design recommendations for*

    adaptive intelligent tutoring systems (Vol. I, Learner Modeling; pp. 165-182). Orlando,

    FL: US Army Research Laboratory.

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).

Springer-Verlag, San Antonio, TX, USA.

Aleven, V., Koedinger, K.R., Sinclair, H. C., & Snyder, J. (1998). Combatting shallow learning

in a tutor for geometry problem solving. In B. P. Goettl, H. M. Halff, C. L. Redfield, &

V. J. Shute (Eds.), *Lecture Notes in Computer Science: Vol 1452. Proceedings of the 4th*

International Conference on Intelligent Tutoring Systems, ITS 1998 (pp. 364-373).

Berlin: Springer Verlag.

Aleven, V., McLaren, B. M., Roll, I., & Koedinger, K. R. (2006). 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., Roll, I., McLaren, B. M., & Koedinger, K. R. (2010). Automated, unobtrusive,

action-by-action assessment of self-regulation during learning with an intelligent tutoring

system. *Educational Psychologist, 45*(4), 224-233.

Aleven, V., Roll, I., McLaren, B. M., & Koedinger, K. R. (2016). Help helps, but only so much:

Research on help seeking with intelligent tutoring systems. *International Journal of*

Artificial Intelligence in Education, 26(1*)*.

Aleven, V., Sewall, J., Popescu, O., Xhakaj, F., Chand, D., Baker, R., ... Gasevic, D. (2015). The

beginning of a beautiful friendship?: Intelligent tutoring systems and MOOCs. In C.

Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), *Artificial Intelligence in*

Education: 17th International Conference, AIED 2015,. (Vol. 9112, pp. 525-528). Berlin:

Springer. doi:10.1007/978-3-319-19773-9_53

Anand, P. G., & Ross, S. M. (1987). Using computer-assisted instruction to personalize

arithmetic materials for elementary school children. *Journal of Educational Psychology,*

79(1), 72-78. doi:10.1037/0022-0663.79.1.72

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., Cooper, D. G., Burleson, W., Woolf, B. P., Muldner, K., & Christopherson, R. (2009). Emotion sensors go to school. In V. Dimitrova, R. Mizoguchi, & B. du Boulay (Eds.), *Proceeding of the 2009 Conference on Artificial Intelligence in Education: Building learning systems that care: From knowledge representation to affective modelling* (pp. 17-24). Amsterdam: IOS Press.

Arroyo, I., Ferguson, K., Johns, J., Dragon, T., Mehranian, H., Fisher, D., Barto, A., Mahadevan, S. & Woolf, B. (2007). Repairing disengagement with non invasive interventions. In R. Luckin, K. Koedinger & J. Greer (Eds), *Proceedings of the 13th International Conference on Artificial Intelligence in Education* (pp. 195–202). Los Angeles: IOS Press.

Arroyo, I., Mehranian, H., & Woolf, B. P. (2010). Effort-based tutoring: An empirical approach to intelligent tutoring. 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. 1-10). Worcester, MA: International Educational Data Mining Society.

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

Atkinson, R. C. (1972). Optimizing the learning of a second-language vocabulary. *Journal of Experimental Psychology, 96*(1)*,* 124-129. http://dx.doi.org/10.1037/h0033475

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), *Intelligent tutoring systems, Volume 7315 Lecture*
>
> *Notes in Computer Science* (pp. 212-221). Springer Berlin Heidelberg. doi:10.1007/978-
>
> 3-642-30950-2_27

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). Springer New York. doi:10.1007/978-1-4419-
>
> 5546-3_7

Baker, R. S., de Carvalho, A., Raspat, J., Aleven, V., Corbett, A. T., & Koedinger, K. R. (2009).

> Educational software features that encourage and discourage "gaming the system". 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. 475-
>
> 482). Amsterdam: IOS Press.

Baker, R, S.J.d., Gowda, S. M., Wixon, M., Kalka, J., Wagner, A. Z., Salvi, A., Aleven, V.,

> Kusbit, G. W., Ocumpaugh, J., & Rossi, L. (2012). Towards sensor-free affect detection
>
> in Cognitive Tutor Algebra. 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. 126-133). Worcester, MA: International Educational Data
>
> Mining Society.

Baker, R., Walonoski, J., Heffernan, N., Roll, I., Corbett, A., & Koedinger, K. (2008). Why

students engage in "gaming the system" behavior in interactive learning environments. *Journal of Interactive Learning Research, 19*(2), 185-224.

Beck, J. E., Woolf, B. P. and Beal, C. R. (2000) ADVISOR: A machine learning architecture for intelligent tutor construction. In H. Kautz & B. Porter (Eds.), *Proceedings of the 17th National Conference on Artificial Intelligence* (pp. 552–557). Menlo Park, CA: The AAAI Press.

Bernacki, M., & Walkington, C. (2014). The Impact of a Personalization Intervention for Mathematics on Learning and Non-Cognitive Factors. In Stamper, J., Pardos, Z., Mavrikis, M., McLaren, B.M. (eds.) *Proceedings of the 7th International Conference on Educational Data Mining*, London.

Bloom, B. S. (1968). Learning for mastery. *Evaluation Comment, 1*(2), 112. Los Angeles: University of California at Los Angeles, Center for the Study of Evaluation of Instructional Programs.

Boekaerts, M. (2007). Understanding students' affective processes in the classroom. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 37-56). San Diego, CA: Academic Press.

Booth, J. L., Barbieri, C., Eyer, F., & Pare-Blagoev, E. J. (2014). Persistent and pernicious errors in algebraic problem solving. *The Journal of Problem Solving, 7*(1), Article 3. Retrieved from http://docs.lib.purdue.edu/jps/vol7/iss1/3

Booth, J.L., Lange, K.E., Koedinger, K.R., & Newton, K.J. (2013). Example problems that improve student learning in algebra: Differentiating between correct and incorrect examples. *Learning and Instruction, 25*, 24-34.

Brown, E. J., Brailsford, T. J., Fisher, T., & Moore, A. (2009). Evaluating learning style

personalization in adaptive systems: quantitative methods and approaches. *IEEE Transactions on Learning Technologies, 2*(1), 10–22.

Brusilovsky, P. (2001). Adaptive Hypermedia. *User Modeling and User-Adapted Interaction*, *11*(1-2), 87-110.

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

Calvo, R. A., & D'Mello, S. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Transactions on Affective Computing, 1*(1), 18-37.

Chi, M., VanLehn, K., Litman, D., & Jordan, P. (2011). Empirically evaluating the application of reinforcement learning to the induction of effective and adaptive pedagogical strategies. *User Modeling and User-Adapted Interaction, 21*(1-2), 137-180. doi:10.1007/s11257-010-9093-1

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,* 145-182.

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

Clark, R.E., Feldon, D., van Merriënboer, J., Yates, K., & Early, S. (2007). Cognitive task analysis. In J.M. Spector, M.D. Merrill, J.J.G. van Merriënboer, & M.P. Driscoll (Eds.), *Handbook of Research on Educational Communications and Technology* (3rd ed., pp. 577–593). Mahwah, NJ: Lawrence Erlbaum Associates.

Coffield, F., Moseley, D., Hall, E., Ecclestone, K., & Vorhaus, D. J. (2004). *Learning styles and*

pedagogy in post-16 learning. A systematic and critical review. London: The Learning and Skills Research Centre.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: Teaching the crafts of reading, writing, and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instructions: Essays in honor of Robert Glaser* (pp. 453-494). Hillsdale, NJ: Lawrence Erlbaum Associates.

Conati, C. (2016). Commentary on: "Toward computer-based support of metacognitive skills: A computational framework to coach self explanation." *International Journal of Artificial Intelligence in Education, 26*(1).

Conati, C. (2013). Modeling and scaffolding self-explanation across domains and activities. In R. Azevedo, & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 367-383). Springer New York. doi:10.1007/978-1-4419-5546-3_24

Conati, C. (2002). Probabilistic assessment of users' emotions in educational games. *Applied Artificial Intelligence, 16*(7-8), 555-575. doi:10.1080/0883951029003039

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., & Maclaren, H. (2005). Data-driven refinement of a probabilistic model of user affect. In L. Ardissono, P. Brna, & A. Mitrovic (Eds.), *Proceedings of the 10th International Conference on User Modeling, UM 2005* (pp. 40-49). Berlin, Heidelberg: Springer. doi:10.1007/11527886_

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*, 398-415.

Constantinidou, F., & Baker, S. (2002). Stimulus modality and verbal learning performance in

    normal aging. *Brain and Language, 82*, 296–311.

Cook, D. A. (2012). Revisiting cognitive and learning styles in computer-assisted instruction:

    Not so useful after all. *Academic Medicine, 87*(6), 778–784. doi:

    10.1097/ACM.0b013e3182541286

Cook, D. A., Thompson, W. G., Thomas, K. G., & Thomas, M. R. (2009). Lack of interaction

    between sensing-intuitive learning styles and problem-first versus information-first

    instruction: a randomized crossover trial. *Advances in Health Sciences Education: Theory*

    and Practice, 14(1), 79–90. http://doi.org/10.1007/s10459-007-9089-8

Corbett, A. T. and Anderson, J. R. (1995). Knowledge tracing: Modeling the acquisition of

    procedural knowledge. *User Modeling and User-Adapted Interaction, 4,* 253–278.

Corbett, A., McLaughlin, M. and Scarpinatto, K. C. (2000). Modeling student knowledge:

    Cognitive tutors in high school and college. *User Modeling and User-Adapted*

    Interaction, *10*, 81-108.

Corbett, A., Wagner, A., & Raspat, J. (2003). The impact of analysing example solutions on

    problem solving in a pre-algebra tutor. In U. Hoppe, F. Verdejo, & J. Kay (Eds.),

    Proceedings of the 11th International Conference Artificial Intelligence in Education,

    AIED 2003 (pp. 133-140). Amsterdam: IOS Press.

Cordova, D. I., & Lepper, M. R. (1996). Intrinsic motivation and the process of learning:

    Beneficial effects of contextualization, personalization, and choice. *Journal of*

    Educational Psychology, 88(4), 715-730.

Cronbach, L. & Snow, R. (1977). *Aptitudes and instructional methods: A handbook for research*

    on interactions. New York: Irvington.

Desmarais, M. C., & Baker, R. S. (2012). A review of recent advances in learner and skill

   modeling in intelligent learning environments. *User Modeling and User-Adapted*

   Interaction, 22(1-2), 9-38.

Despotovic-Zrakic, M., Markovic, A., Bogdanovic, Z., Barac, D., & Krco, S. (2012). Providing

   Adaptivity in Moodle LMS courses. *Journal of Educational Technology & Society, 15*(1),

   326-338.

D'Mello, S. (2013). A selective meta-analysis on the relative incidence of discrete affective

   states during learning with technology. *Journal of Educational Psychology, 105(4),* 1082-

   1099.

D'Mello, S., Blanchard, N., Baker, R., Ocumpaugh, J., & Brawner, K. (2014). I feel your pain:

   A selective review of affect-sensitive instructional strategies. In R. Sottilare, A. Graesser,

   X. Hu, & H. Holden (Eds.), Design recommendations for adaptive intelligent tutoring

   systems learner modeling (Volume II, Instructional Management, pp. 169-182). Orlando,

   FL: US Army Research Laboratory.

D'Mello, S. K., & Graesser, A. C. (2014). Feeling, thinking, and computing with affect-aware

   learning. In R. A. Calvo, S. K. D'Mello, J. Gratch, & A. Kappas (Eds.), *The Oxford*

   handbook of affective computing (pp. 419-434). Oxford University Press. doi:

   10.1093/oxfordhb/9780199942237.013.032.

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for

   learning. *Learning and Instruction, 29,* 153-170. doi:10.1016/j.learninstruc.2012.05.003

D'Mello, S., Lehman, B., Sullins, J., Daigle, R., Combs, R., Vogt, K., Perkins, L., & Graesser, A.

   (2010). A time for emoting: When affect-sensitivity is and isn't effective at promoting

   deep learning. In J. Kay & V. Aleven (Eds.), *Proceedings of the 10th International*

Conference on Intelligent Tutoring Systems (pp. 245-254). Berlin / Heidelberg: Springer.

D'Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal Human-Computer Studies, 70*(5), 377-398. doi:10.1016/j.ijhcs.2012.01.004

Dochy, F., Segers, M. & Pletinckx, J. (2002). The question of entry assessment or how can we assess previously acquired knowledge? Main research findings and implications for practice. *Journal of Continuing Engineering Education and Life*, *12*, 31-44.

Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior, 52,* 338-348. doi:10.1016/j.chb.2015.05.041

Ericsson, K. A., & Simon, H. A. (1984). *Protocol Analysis: Verbal Reports as Data*. Cambridge, MA: The MIT Press.

Falmagne, J. -C., Albert, D., Doble, C., Eppstein, D., & Hu, X. (2013). *Knowledge spaces: Applications in education.* Springer Science & Business Media.

Falmagne, J. -C., Koppen, M., Villano, M., Doignon, J. -P., & Johannesen, L. (1990). Introduction to knowledge spaces: How to build, test and search them. *Psychological Review, 97*(2), 201-224.

Felder, R. M., & Silverman, L. K. (1988). Learning and teaching styles in engineering education. *Engineering Education, 78*(7), 674–681.

Felder, R. M., & Soloman, B. A. (1997). *Index of Learning Styles questionnaire.* Retrieved Jan 15, 2016, from http://www.igi-global.com/pdf.aspx?tid=121101&ptid=981&ctid=17&t=appendix%20a

Forbes-Riley, K., & Litman, D. (2011). Benefits and challenges of real-time uncertainty

detection and adaptation in a spoken dialogue computer tutor. *Speech Communication*, *53*(9-10), 1115-1136. doi:10.1016/j.specom.2011.02.006

Ford, N. (1985). Learning styles and strategies of postgraduate students. *British Journal of Educational Technology, 16*(1), 65-79.

Ford, N., & Chen, S. Y. (2001). Matching/mismatching revisited: An empirical study of learning and teaching styles. *British Journal of Educational Technology, 32*(1), 5–22.

Goguadze, G., Sosnovsky, S., Isotani, S., & McLaren, B.M. (2011). Evaluating a Bayesian student model of decimal misconceptions. In M. Pechenizkiy, T. Calders, C. Conati, S. Ventura, C. Romero & John Stamper (Eds.), *Proceedings of the 4th International Conference on Educational Data Mining (EDM 2011)* (pp. 301-306). Worcester, MA: International Educational Data Mining Society.

Grawemeyer, B., Mavrikis, M., Holmes, W., & Gutierrez-Santos, S. (2015). Adapting Feedback Types According to Students' Affective States. In C. Conati, N. Heffernan, A. Mitrovic, M.F. Verdejo (Eds), *Proceedings of the 17th International Conference on Artificial Intelligence in Education, AIED 2015* (pp. 586-590). New York: Springer.

Graf, S., & Kinshuk. (2007). Providing adaptive courses in learning management systems with respect to learning styles. In T. Bastiaens, & S. Carliner (Eds.), *Proceedings of World Conference on e-learning in Corporate, Government, Healthcare, and Higher Education* (pp. 2576–2583). Chesapeake, VA: AACE.

Harackiewicz, J. M., Durik, A. M., Barron, K. E., Linnenbrink-Garcia, L., & Tauer, J. M. (2008). The role of achievement goals in the development of interest: Reciprocal relations between achievement goals, interest, and performance. *Journal of Educational Psychology, 100(1),* 105-122. doi:10.1037/0022-0663.100.1.105

Harley, J. M., Bouchet, F., Hussain, M. S., Azevedo, R., & Calvo, R. (2015). A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior, 48,* 615-625. doi:0.1016/j.chb.2015.02.013

Hausmann, R. G. M., Nokes, T. J., VanLehn, K., & Gershman, S. (2009). The design of self-explanation prompts: The fit hypothesis. In N. Taatgern & H. van Rijn (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 2626–2631). Austin, TX: Cognitive Science Society.

Heffernan, N. & Koedinger, K. R. (1997). The composition effect in symbolizing: The role of symbol production vs. text comprehension. In Shafto, M. G. & Langley, P. (Eds.) *Proceedings of the Nineteenth Annual Conference of the Cognitive Science Society*, (pp. 307-312). Hillsdale, NJ: Erlbaum.

Heilman, M., Collins-Thompson, K., Callan, J., Eskenazi, M., Juffs, A., & Wilson, L. (2010). Personalization of reading passages improves vocabulary acquisition. *International Journal of Artificial Intelligence in Education, 20*(1), 73-98.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127.

Howard-Jones, P. A. (2014). Neuroscience and education: myths and messages. *Nature Reviews Neuroscience, 15*(12), 817–824. http://doi.org/10.1038/nrn38

Hulleman, C. S., & Harackiewicz, J. M. (2009). Promoting interest and performance in high school science classes. *Science, 326*(5958), 1410-1412. doi:10.1126/science.1177067

Isotani, S., McLaren, B.M., & Altman, M. (2010). Towards intelligent tutoring with erroneous examples: A taxonomy of decimal misconceptions. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems*

(ITS-2010). *Lecture Notes in Computer Science,* 6094 (pp. 346-348). Berlin: Springer.

Jonassen, D. H., & Grabowski, B. L. (1993). *Handbook of individual differences, learning, and instruction*. Hillsdale, NJ: Erlbaum.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*(4), 509-539. doi:10.1007/s10648-007-9054-3

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). The expertise reversal effect. *Educational Psychologist, 38*(1), 23-31. doi:10.1207/S15326985EP3801_4

Kalyuga, S., & Sweller, J. (2005). Rapid dynamic assessment of expertise to improve the efficiency of adaptive e-learning. *Educational Technology Research and Development, 53*(3), 83–93. http://doi.org/10.1007/BF02504800

Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction, *Journal of Educational Psychology, 96*(3), 558-568, http://dx.doi.org/10.1037/0022-0663.96.3.558

Keller, F. S. (1968). "Good-bye, teacher...". *Journal of Applied Behavior Analysis, 1*(1), 79-89. doi:10.1901/jaba.1968.1-79

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

Kirschner, P. A., & van Merriënboer, J. J. (2013). Do learners really know best? Urban legends in education. *Educational Psychologist, 48*(3), 169–183. doi:10.1080/00461520.2013.804395

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

Koedinger, K. R., & Anderson, J. R. (1998). Illustrating principled design: The early evolution of

a cognitive tutor for algebra symbolization. *Interactive Learning Environments, 5*, 161-

180.

Koedinger, K. R., Brunskill, E., S.J.d. Baker, R., McLaughlin, E. A., & Stamper, J. C. (2013a).

New potentials for data-driven intelligent tutoring system development and optimization.

AI Magazine, 34(3), 27-41.

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., Corbett, A.C., & Perfetti, C. (2012). The Knowledge-Learning-Instruction

(KLI) framework: Bridging the science-practice chasm to enhance robust student

learning. *Cognitive Science, 36* (5)*, 757-798. ISSN: 0364-0213 print / 1551-6709 online

DOI: 10.1111/j.1551-6709.2012.01245.x

Koedinger, K.R. & McLaughlin, E.A. (2010). Seeing language learning inside the math:

Cognitive analysis yields transfer. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of

the 32nd Annual Conference of the Cognitive Science Society*. (pp. 471-476.) Austin, TX:

Cognitive Science Society.

Koedinger, K.R. & Nathan, M.J. (2004). The real story behind story problems: Effects of

representations on quantitative reasoning. *The Journal of the Learning Sciences, 13*(2),

129-164.

Koedinger, K. R., Pavlik, P., McLaren, B. M., & Aleven, V. (2008). Is it better to give than to

receive? The assistance dilemma as a fundamental unsolved problem in the cognitive

science of learning and instruction. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.),

Proceedings of the 30th annual meeting of the Cognitive Science Society (pp. 2155-

2160). Austin, TX: Cognitive Science Society.

Koedinger, K. R., Stamper, J. C., McLaughlin, E. A., & Nixon, T. (2013b). Using data-driven

discovery of better student models to improve student learning. In H.C. Lane, K. Yacef,

J. Mostow, P. Pavlik, *Proceedings of the 16th International Conference on Artificial

Intelligence in Education (*pp. 421-430). Berlin Heidelberg: Springer.

Kulik, C., Kulik, J.A., & Bangert-Drowns, R.L. (1990). Effectiveness of mastery learning

programs: A meta-analysis. *Review of Educational Research, 60*(2), 265-299.

Lee, S. Y., Rowe, J. P., Mott, B. W., & Lester, J. C. (2014). A supervised learning framework for

modeling director agent strategies in educational interactive narrative. *IEEE Transactions

on Computational Intelligence and AI in Games, 6*(2), 203-215.

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.

Lehman, B., Matthews, M., D'Mello, S., & Person, N. (2008). What are you feeling?

Investigating student affective states during expert human tutoring sessions. In B. Woolf,

E. Aimeur, R. Nkambou, & S. Lajoie (Eds.), *Proceedings of the 9th International

Conference on Intelligent Tutoring Systems, ITS 2008* (pp. 50-59). Berlin, Heidelberg:

Springer-Verlag. doi:10.1007/978-3-540-69132-7_10

Lehman, B., D'Mello, S. K., Strain, A. C., Mills, C., Gross, M., Dobbins, A., . . . Graesser, A. C.

(2013). Inducing and tracking confusion with contradictions during complex learning.

International Journal of Artificial Intelligence in Education, *22*(1-2), 85-105.

Limongelli, C., Sciarrone, F., Temperini, M., & Vaste, G. (2009). Adaptive learning with the LS-

plan system: A field evaluation. *IEEE Transactions on Learning Technologies, 2*(3),

203-215.

Long, Y. (2015). *Supporting learner-controlled problem selection in intelligent tutoring systems.* (Unpublished doctoral dissertation). Human-Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA.

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., Meyer, O., Thille, C. (2008). The Open Learning Initiative: Measuring the effectiveness of the OLI learning course in accelerating student learning. *Journal of Interactive Media in Education*, 1, Art-13. V. http://jime.open.ac.uk/article/view/2008-14/351.

Mabbott, A., & Bull, S. (2004). Alternative views on knowledge: Presentation of open learner models. In J. C. Lester, R. M. Vicario, & F. Paraguaçu (Eds.), *Proceedings of Seventh International Conference on Intelligent Tutoring Systems, ITS 2004* (pp. 689-698). Springer Berlin Heidelberg. doi:10.1007/978-3-540-30139-4_65

Mampadi, F., Chen, S. Y., Ghinea, G., & Chen, M.-P. (2011). Design of adaptive hypermedia learning systems: A cognitive style approach. *Computers & Education, 56(4),* 1003–

1011. http://doi.org/10.1016/j.compedu.2010.11.018

Massa, L., & Mayer, R. E. (2006). Testing the ATI hypothesis: Should multimedia instruction accommodate verbalizer-visualizer cognitive style? *Learning and Individual Differences*, *16*(4), 321–335. http://doi.org/10.1016/j.lindif.2006.10.001

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

Mazziotti, C., Holmes, W., Wiedmann, M., Loibl, K., Rummel, N., Mavrikis, M., Hansen, A., Grawemeyer, B. (2015). Robust Student Knowledge: Adapting to Individual Student Needs as They Explore the Concepts and Practice the Procedures of Fractions. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), *Artificial Intelligence in Education: 17th International Conference, AIED 2015: 4th Workshop on Intelligent Support for Learning in Groups (ISLG)* (pp. 32-40). Berlin: Springer.

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., Adams, D., Durkin, K., Goguadze, G., Mayer, R.E., Rittle-Johnson, B., Sosnovsky, S., Isotani, S. & van Velsen, M. (2012). To err is human, to explain and correct is divine: A study of interactive erroneous examples with middle school math students. In Ravenscroft, A., Lindstaedt, S., Delgado Kloos, C. and Hernándex-Leo, D. (Eds): *Proceedings of ECTEL 2012: Seventh European Conference on Technology Enhanced Learning*, LNCS 7563 (pp.222–235). Springer, Berlin,.

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

McNamara, D. S., O'Reilly, T., Rowe, M., Boonthum, C., & Levinstein, I. B. (2007). ISTART: A web-based tutor that teaches self-explanation and metacognitive reading strategies. In D. S. McNamara (Ed.), *Reading comprehension strategies: Theories, interventions, and technologies* (pp. 397-421). Erlbaum. Mahwah, NJ.

McQuiggan, S. W., Mott, B. W., & Lester, J. C. (2008). Modeling self-efficacy in intelligent tutoring systems: An inductive approach. *User Modeling and User-Adapted Interaction, 18*(1), 81-123.

Meluso, A., Zheng, M., Spires, H. A., & Lester, J. (2012). Enhancing 5th graders' science content knowledge and self-efficacy through game-based learning. *Computers and Education, 59(2)*, 497-504.

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.

Mitrovic, A., Ohlsson, S., Barrow, DK. (2013). The effect of positive feedback in a constraint-based intelligent tutoring system. *Computers & Education*, *60* (1), 264-272.

Muldner, K., & Conati, C. (2007). Evaluating a decision-theoretic approach to tailored example selection. In M. Veloso (Ed.), *Proceedings of the 20th International Joint Conference on Artificial Intelligence*, *IJCAI 2007* (pp. 483-488). San Francisco, CA: Morgan Kaufmann.

Murray, R.C., VanLehn, K. & Mostow, J. (2004). Looking ahead to select tutorial actions: A decision-theoretic approach. *International Journal of Artificial Intelligence in Education, 14*(3-4), 253-278.

Mustafa, Y. E. A., & Sharif, S. M. (2011). An approach to adaptive e-learning hypermedia system based on learning styles (AEHS-LS): Implementation and evaluation.

International Journal of Library and Information Science, 3(1), 15–28.

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

Nückles, M., Hübner, S., Dümer, S., & Renkl, A. (2010). Expertise reversal effects in writing-to-learn. *Instructional Science, 38*(3), pp 237-258.

Oksa, A., Kalyuga , S., & Chandler, P. (2010). Expertise reversal effect in using explanatory notes for readers of Shakespearean text. *Instructional Science, 38*(3), pp 217-236.

Özyurt, Ö., & Özyurt, H. (2015). Learning style based individualized adaptive e-learning environments: Content analysis of the articles published from 2005 to 2014. *Computers in Human Behavior, 52,* 349–358. doi:10.1016/j.chb.2015.06.020

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9(3)*, 105–119. doi:10.1111/j.1539-6053.2009.01038.x

Pask, G. (1976). Styles and strategies of learning. *British Journal of Educational Psychology, 46*, 128–148.

Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). Control-value theory of achievement emotions: An integrative approach to emotions in education. In P. A. Schutz & R. Pekrun (Eds.), *International Handbook of Emotions in Education* (pp. 13-36). Amsterdam: Academic Press.

Piaget, J. (1952). *The origins of intelligence*. New York: International University Press.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology*, *82*(1), 33-40.

Popescu, E. (2010). Adaptation provisioning with respect to learning styles in a Web-based

educational system: an experimental study. *Journal of Computer Assisted Learning, 26*(4), 243-257.

Popescu, E. (2009). Evaluating the Impact of Adaptation to Learning Styles in a Web-Based Educational System. In M. Spaniol, Q. Li, R. Klamma, & R. W. H. Lau (Eds.), *Advances in Web Based Learning – ICWL 2009* (pp. 343–352). Springer Berlin Heidelberg. Retrieved from http://link.springer.com/chapter/10.1007/978-3-642-03426-8_41

Rau, M. A., Aleven, V., & Rummel, N. (2009). Intelligent tutoring systems with multiple representations and self-explanation prompts support learning of fractions`. 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. 441-448). Amsterdam: IOS Press.

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.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99*(3), 561-574. doi:10.1037/0022-0663.99.3.561

Rohrer, D., & Pashler, H. (2012). Learning styles: where's the evidence? *Medical Education, 46*(7), 634–635. http://doi.org/10.1111/j.1365-2923.2012.04273.x

Roll, I., Aleven, V., & Koedinger, K. R. (2010). The invention lab: Using a hybrid of model tracing and constraint- based modeling to offer intelligent support in inquiry environments. In V. Aleven, J. Kay, & J. Mostow (Eds.), *Proceedings of the 10th International Conference on Intelligent Tutoring Systems* (pp. 115-24). Berlin: Springer

Verlag.

Roll, I., Aleven, V., McLaren, B. M., & Koedinger, K. R. (2011). Improving students' help-seeking skills using metacognitive feedback in an intelligent tutoring system. *Learning and Instruction, 21*(2*)*, 267-280.

Roll, I., Stampfer Wiese, E., Long, Y., Aleven, V., & Koedinger, K. R. (2014). 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.

Rowe, J. P., & Lester, J. C. (2015). Improving student problem solving in narrative-centered learning environments: A modular reinforcement learning framework. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), *Proceedings of the 17th international conference on artificial intelligence in education, AIED 2015* (pp. 419-428). New York: Springer International Publishing. doi:10.1007/978-3-319-19773-9_42

Sabourin, J., Mott, B., & Lester, J. C. (2011). Modeling learner affect with theoretically grounded dynamic bayesian networks. In S. D'Mello, A. Graesser, B. Schuller, & J. C. Martin (Eds.), *Proceedings 4th international conference, on affective computing and intelligent interaction, ACII 2011* (pp. 286-295). Springer Berlin Heidelberg. doi:10.1007/978-3-642-24600-5_3

Salden, R. J. C. M., Aleven, V., Schwonke, R., & Renkl, A. (2010). The expertise reversal effect and worked examples in tutored problem solving. *Instructional Science, 38(*3*),* 289-307. doi:10.1007/s11251-009-9107-8

Salden, R. J., Paas, F., & van Merriënboer, J. J. (2006a). A comparison of approaches to learning

task selection in the training of complex cognitive skills. *Computers in Human Behavior*, *22*(3), 321-333. doi:10.1016/j.chb.2004.06.003

Salden, R. J., Paas, F., & van Merriënboer, J. J. (2006b). Personalised adaptive task selection in air traffic control: Effects on training efficiency and transfer. *Learning and Instruction*, *16*(4), 350-362. doi:10.1016/j.learninstruc.2006.07.00

Schunk, D. H., Pintrich, P. R., & Meece, J. L. (2008). *Motivation in education: Theory, research, and applications*. Upper Saddle River, NJ: Pearson/Merrill Prentice Hall.

Snow, R. (1989). Aptitude-Treatment Interaction as a framework for research on individual differences in learning. In P. Ackerman, R.J. Sternberg, & R. Glaser (ed.), *Learning and Individual Differences*. New York: W.H. Freeman.

Sottilare, R., Graesser, A., Hu, X., & Holden, H. (2013). Design recommendations for adaptive intelligent tutoring systems (Vol. I, Learner Modeling). Orlando, FL: US Army Research Laboratory.

Stamper, J., Eagle, M., Barnes, T., & Croy, M. (2013). Experimental evaluation of automatic hint generation for a logic tutor. *International Journal of Artificial Intelligence in Education,* *22*(1-2), 3-17. doi:10.3233/JAI-130029

Stamper, J. & Koedinger, K.R. (2011). Human-machine student model discovery and improvement using data. In G. Biswas, S. Bull, J. Kay, A. Mitrovic (Eds.), *Proceedings of the 15th International Conference on Artificial Intelligence in Education (353-360)*. Berlin, Heidelberg: Springer.

Stampfer, E., & Koedinger, K. R. (2013). Conceptual scaffolding to check one's procedures. 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. 916-919).

Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-39112-5_145

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.), *Proceedings of the 16th international conference on artificial intelligence in education, AIED 2013* (pp. 239-248). Berlin, Heidelberg: Springer. doi:10.1007/978-3-642-39112-5_25

Tan, J., Biswas, G. & Schwartz, D. (2006). Feedback for metacognitive support in learning by teaching environments. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual meeting of the Cognitive Science Society* (pp. 828-833). Mahwah, NJ: Lawrence Erlbaum.

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research, 64*(1*)*, 37-54. doi:10.3102/00346543064001037

Tseng, J. C. R., Chu, H. C., Hwang, G. J., & Tsai, C. C. (2008). Development of an adaptive learning system with two sources of personalization information. *Computers & Education, 51*(2), 776–786.

Vail, A. K., Boyer, K. E., Wiebe, E. N., & Lester, J. C. (2015). The mars and venus effect: The influence of user gender on the effectiveness of adaptive task support. In F. Ricci, K. Bontcheva, O. Conlan, & S. Lawless (Eds.), *Proceedings of the 23rd international conference on user modeling, adaptation and personalization, UMAP 2015* (pp. 265-276). Springer International Publishing. doi:10.1007/978-3-319-20267-9_22

Vandewaetere, M., & Clarebout, G. (2014). Advanced technologies for personalized learning, instruction, and performance. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp.

425-437). New York: Springer. doi:10.1007/978-1-4614-3185-5_34

Vandewaetere, M., Desmet, P., & Clarebout, G. (2011). The contribution of learner

characteristics in the development of computer-based adaptive learning environments.

Computers in Human Behavior, 27(1), 118-130. doi:10.1016/j.chb.2010.07.038

VanLehn, K. (2016). Regulative loops, step loops and task loops. *International Journal of

Artificial Intelligence in Education, 26*(1).

VanLehn, K. (2011). The Relative Effectiveness of Human Tutoring, Intelligent Tutoring

Systems, and Other Tutoring Systems, *Educational Psychologist, 46*(4), 197-221, DOI:

10.1080/00461520.2011.611369

VanLehn, K. (2006). The behavior of tutoring systems. *International Journal of Artificial

Intelligence in Education. 16*(3), 227-265.

VanLehn, K., Freedman, R., Jordan, P., Murray, C., Osan, R., Ringenberg, M., . . . Wintersgill,

M. (2000). Fading and deepening: The next steps for andes and other model-tracing

tutors. In G. Gauthier, C. Frasson, & K. VanLehn (Eds.), *Proceedings of the 5th

International Conference on Intelligent Tutoring Systems, ITS 2000* (pp. 474-483).

Springer Berlin Heidelberg. doi:10.1007/3-540-45108-0_51

van Merriënboer, J. J. G., Salden, R., Corbalan, G., de Croock, M., Kester, L. & Paas, F. (2004).

Dynamic selection of learning tasks according to the 4C/ID-Model. *Association for

Educational Communications and Technology* (pp. 640-647).

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*.

Cambridge, MA: Harvard University Press.

Waalkens, M., Aleven, V., & Taatgen, N. (2013). Does supporting multiple student strategies

lead to greater learning and motivation? Investigating a source of complexity in the

architecture of intelligent tutoring systems. *Computers & Education*, *60*(1), 159 - 171.

doi:10.1016/j.compedu.2012.07.016

Wagster, J., Tan, J., Wu, Y., Biswas, G., & Schwartz, D. (2007). Do learning by teaching

environments with metacognitive support help students develop better learning

behaviors? In D. S. McNamara & J. G. Trafton (Eds.), *Proceedings of the 29th Meeting*

of the Cognitive Science Society (pp. 695-700). Austin, TX: Cognitive Science Society.

Walkington, C. (2013). Using learning technologies to personalize instruction to student

interests: The impact of relevant contexts on performance and learning outcomes. *Journal*

of Educational Psychology, 105(4), 932-945.

Walkington, C. & Bernacki, M. (2015). *The Effects of Personalization of Algebra Instruction to*

Students' Interests on Learning, Behavior, and Interest in Mathematics. Presentation at

2015 Annual Meeting of the American Educational Research Association, Chicago, IL.

Walkington, C., & Sherman, M. (2012). Using adaptive learning technologies to personalize

instruction: The impact of interest-based scenarios on performance in algebra. In van

Aalst, J., Thompson, K., Jacobson, M., & Reimann, P. (Eds.), *Proceedings of the 10th*

International Conference of the Learning Sciences (pp. 932-945). Sydney, NSW,

Australia.

Weerasinghe, A., & Mitrovic, A. (2006). Facilitating deep learning through self-explanation in

an open-ended domain. *International Journal of Knowledge-based and Intelligent*

Engineering Systems, 10(1), 3-19.

Willingham, D. T., Hughes, E. M., & Dobolyi, D. G. (2015). The Scientific Status of Learning

Styles Theories. *Teaching of Psychology, 42*(3), 266–271.

http://doi.org/10.1177/0098628315589505

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, NJ: Erlbaum.

Witkin, H. A., Oltman, P. K., Raskin, I., & Karp, S. A. (1971). *A manual for the Embedded-Figures Test*. Palo Alto: Consulting Psychologists Press.

Wylie, R., Sheng, M., Mitamura, T., & Koedinger, K. (2011). Effects of adaptive prompted self-explanation on robust learning of second language grammar. In G. Biswas, S. Bull, J. Kay, & A. Mitrovic (Eds.), *Lecture Notes in Computer Science: Artificial Intelligence in Education* (Vol. 6738, pp. 588-590). Springer Berlin Heidelberg. doi:10.1007/978-3-642-21869-9_110

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

Zimmerman, B. J., & Campillo, M. (2003). Motivating self-regulated problem solvers. In J. E. Davidson & R. E. Sternberg (Eds.), *The Psychology of Problem Solving* (pp. 233- 262). Cambridge, UK: Cambridge University Press.

Zimmerman, B. J. & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology 80*(3). doi:10.1037/0022-0663.80.3.284.

Zimmerman, B. J. & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. J*ournal of Educational Psychology*, *82*(1), 51-59.