Abstract
Good teachers are especially effective at providing instruction when they are able to make a connection with a student that involves both cognitive and non-cognitive factors. We explore research into both cognitive and non-cognitive personalization of intelligent tutoring system (ITS) content for mathematics to provide experiences more like those of one-on-one instruction. We review Carnegie Learning’s approach, in its Cognitive Tutor and MATHia X ITSs, to competency-based progression as a means of adapting or personalizing mathematics instruction and practice based on cognitive factors (i.e., skill mastery) and then explore recent research on personalizing mathematics word problems based on student interest areas as a means of personalization based on non-cognitive factors.

Introduction
Good teachers are especially effective at providing instruction when they are able to make a connection with a student that involves both cognitive and non-cognitive factors. Such teachers may be able to provide students with examples and illustrations that address the particular cognitive gaps and misconceptions held by the students and also that provide a real-world context that helps the subject come alive for the student. Bloom (1984) determined that one-to-one human tutoring with mastery learning (Bloom 1968) could improve average student performance by two standard deviations over conventional instruction, a large improvement over conventional instruction. However, providing one-on-one, differentiated instruction and personalized content to individual students in real K-12 classrooms is both expensive and often difficult; class sizes can be large, and instructional time may be cut short, among other practical difficulties.

A major line of artificial intelligence, human-computer interaction, and educational technology research over the past 25+ years has sought to develop intelligent systems that are as effective as one-on-one human tutors in a wide variety of disciplines. One especially successful example of a technology arising out of this tradition is Carnegie Learning’s Cognitive Tutor intelligent tutoring system (ITS) for mathematics (Ritter, et al., 2007). The bedrock of Cognitive Tutor’s success is individualized instruction based on cognitive factors. Cognitive Tutor adaptively presents mathematics problems based on students’ mastery of fine-grained mathematics skills, and progression through content is
largely governed by the pace at which students master skills, echoing as closely as possible the conditions leading to success in Bloom’s study. Although such tutors approach the effectiveness of human tutors (VanLehn, 2011), they do not have the advantage that human tutors have of being able to assess and adapt to students’ interests. We present recent research that experimentally explores the effects on student performance of personalizing the content of mathematics word problems based on student interest areas outside of academic subjects. While adaptive learning technologies like intelligent tutoring systems offer the promise of helping teachers provide personalized instruction to students, such technologies must (attempt to) connect with students by deploying strategies that rely on both cognitive and non-cognitive factors.

**Intelligent Tutoring Systems & Cognitive Tutor**

Carnegie Learning’s ITSs for mathematics, including its Cognitive Tutor and MATHia X products, are based on the ACT-R theory of cognition developed by John R. Anderson and colleagues at Carnegie Mellon University (Anderson, 2007). Cognitive Tutor and MATHia X are parts of a broader blended learning curricular solution provided by Carnegie Learning for Grades 6-8, Algebra I, Geometry, and Algebra II. Student learning in Cognitive Tutor and MATHia X is driven by competency-based progression (also known as mastery learning). Each mathematics topic (i.e., a “section” of content within the software) is associated with a set of fine-grained mathematics skills (or knowledge components) the mastery of which is required for the student to demonstrate competency and progress to the following section of content. Students are presented with complex, multi-step problems (cf. Figure 1), and each of these problems within a section is associated with a subset of skills for the section in which it appears. Steps within each problem provide practice and opportunities for students to demonstrate competency on these fine-grained skills. Problems are assigned to students based on a match between the skills required to solve the problem and the skills within a section that student has not yet mastered.

Student skill mastery is assessed using an artificial intelligence framework called Bayesian Knowledge Tracing (BKT) (Corbett & Anderson 1995) that estimates the probability that a student has mastered a particular skill based on their observed performance in practicing that particular skill in their work with the ITS (i.e., their correct or incorrect answers at steps within a problem). Since each multi-step problem can be solved in different ways, Cognitive Tutor also uses a technique called “model tracing” to track individual student problem-solving strategies; the ITS provides context-sensitive hints and feedback to students, especially by responding to “known misconceptions” in mathematics problem-solving.

**Effectiveness & Implementation**

Recently, the RAND Corporation conducted a large-scale, two-year, randomized field trial of the Cognitive Tutor Algebra curriculum across a set of 147 middle schools and high schools, including over 19,000 students in seven regions of the country (Pane, et al. 2014). To date, this is among the largest, most rigorous effectiveness studies of an educational technology, especially for K-12 mathematics, of its kind. The high school study found a significant positive effect in its second year, the extent of which is
equivalent to an entire year’s additional learning gain in mathematics for the experimental group compared to the control group. This study and several previous, smaller studies (Ritter et al, 2007b; Sarkis, 2004), demonstrate that ITSs like Cognitive Tutor, designed for blended learning, founded on cognitive science principles, and implementing competency-based progression, can effectively enhance student learning outcomes.

Figure 1. This is a screenshot of problem solving in Carnegie Learning’s MATHia X ITS for middle school mathematics. The student is presented with a multi-step word problem and provided with a graphing tool to help provide answers to questions. The green progress bar near the top-right displays student progress toward skills addressed by the section in which this problem is found.

On-going research delves into what drives the effectiveness of solutions, how effective solutions can be deployed even in the first year of implementation, and in which situations ITSs like Cognitive Tutor and MATHia X are most effective (Sales and Pane 2015). Recent work, for example, explores the extent to which teachers in classrooms and computer labs implement competency-based progression or mastery learning in the Cognitive Tutor and how this relates to measures of cognitive factors like student error rates (Ritter, et al. 2016). This work considers how teachers “move along” students in the ITS before they have been judged to have mastered the necessary skills in a particular section to operationalize the notion of deviation from competency-based progression or mastery learning. Teachers move students along for a variety of reasons, including, but not limited to, trying to keep students moving along through material at a teacher-set pace or the notion that, when time is running short in the academic year, perhaps mere exposure to some material is better than working carefully through material until mastery when high-stakes standardized tests are looming.
Nevertheless, deviating from the competency-based progression is likely to produce gaps in student knowledge of pre-requisite skills necessary to make successful progress on skills in later sections. Considering student error rates over time as they progress through material can provide insight into whether such an explanation is plausible. Ritter, et al. (2007) posited that students progressing at their own pace through a well-designed curriculum and system that implements mastery learning ought to have similar error rates over time, roughly corresponding to their general aptitude for mathematics and perhaps other factors about how they interact with an ITS. That is, a well designed, adaptive ITS will present problems and material to students that is appropriate to their present level of knowledge. However, if students are “moved along” or allowed to skip various of pre-requisite skills in earlier sections of course content, when they encounter post-requisite skills and content they will begin to perform worse, resulting in increasing errors rates (as well as greater variability in error rates) over time. Ritter, et al. (2016) found exactly this sort of pattern over the course of a school year at a large school district in the United States with high variability for a variety of characteristics of classroom implementations of the ITS. Students (and classes) that largely were allowed to progress at their own pace experienced relatively constant error rates over the course of the entire academic year versus student who were more frequently moved along in the course material, who experienced steadily increasing error rates over time and greater variability in error rates as the year (and like gaps in students’ pre-requisite knowledge) progressed.

![Figure 2. Scatterplot of time vs. class-level error rate over the course of an academic year, reported in Ritter, et al. (2016). Points in red are for classes that, with relatively high frequency, deviate from mastery learning or competency-based progress. Points in blue represent classes that do so with less frequency.](image)

**Non-Cognitive Factors & Personalization**

While ITSs like Cognitive Tutor can adapt to students’ cognitive needs, researchers are only beginning to explore ways in which they take account of non-cognitive and meta-cognitive factors that may impact learning, including learners’ goals, mood, affect, and interests, possibly leading to even greater effectiveness. One particular non-cognitive
factor around which adaptive elements of ITSs can be developed involves learner interest areas, which are relatively easily ascertainable (i.e., measureable) via survey questions. Just as a human tutor tailors problems and explanations to areas of interest to tutees (e.g., sports), ITSs can present problems and other content personalized to learners’ interests. Within the Cognitive Tutor, a natural place to personalize learning content is within Algebra word problems. Recent experimental work of Walkington and colleagues (Walkington 2013; Walkington and Sherman 2012) explores the impact of such personalization of mathematics story problems on learner performance. Walkington and Sherman (2012) provide examples of an algebra story problem scenario, upon which multi-step problems are based, and personalized versions of the scenario tailored to sports and art interest areas:

- **Original**: One method for estimating the cost of new home construction is based on the proposed square footage of the home. Locally, the average cost per square foot is estimated to be $46.50.
- **Sports**: You are working at the ticket office for a college football team. Each ticket to the first home football game costs $46.50.
- **Art**: You have been working for the school yearbook, taking pictures and designing pages, and now it’s time for the school to sell the yearbooks for $46.50 each.

Early analysis of a randomized controlled trial experiment by Walkington (2013) using these types of problems established that personalized word problems improved student performance, efficiency, and learning gain, especially for algebraic symbolization in a sample of 145 9th grade Algebra I students using Cognitive Tutor Algebra for instruction. Students were randomly assigned to a control condition or an experimental condition in which personalized word problems, in nine categories for which students were asked to rate their interest, were presented in a particular Cognitive Tutor Algebra unit containing problems about linear functions; control condition students in this unit receive standard story problems.

The mechanism by which personalization works has yet to be explored fully. One possibility is that, by presenting problems in the student’s interest area, students understand the relevance of mathematics to their own interests and thus become more interested in mathematics. We refer to this as the “domain interest” hypothesis. Another possibility is that word problems in the student’s interest area are easier for students to understand, since the situations and vocabulary are more familiar to them. We refer to this as the “comprehension” hypothesis. A third explanation is that students are responding in a more general way to the fact that the system is responding to their specification of interest areas. For example, students may come to trust the ITS more or to believe that the ITS understands them personally, because it is using information about their outside interests. We refer to this as the “personal connection” hypothesis.

Creating well-controlled word problems, such as those used by Walkington (2013) results in problems that do not require deep knowledge of the interest area. For example, the sports problem illustrated above does not require any knowledge of football (other than
the fact that tickets are sold to football games), which seems to argue against the “comprehension” hypothesis. A later randomized experiment (Bernacki and Walkington 2014), over a larger number of Cognitive Tutor units and a sample of 154 9th grade Algebra I students, found that word problems more “deeply” personalized to better reflect how students might actually use linear functions in a way that corresponds to their out-of-school interests improve learning compared to problems with only “surface” (or superficial) personalization, but only for those students with lower initial interest in mathematics. These results, coupled with survey results showing that students in the experimental conditions reported statistically significant higher levels of interest in mathematics, tend to support the domain interest hypothesis rather than the comprehension hypothesis.

However, deeper analyses of Walkington’s (2013) experimental data in Walkington and Sherman (2012), address the comprehension hypothesis in a different way, finding that considering the readability level of word problems is especially important. First, their analysis finds a significant increase in performance, controlling for readability of word problems, on both easy and hard knowledge components (as opposed to medium difficulty knowledge components). They also find statistically significant interactions between readability level of word problems and condition (experimental versus control) and between readability level and knowledge component difficulty. For both easy and hard knowledge components, students presented personalized word problems performed significantly better than those presented normal, non-personalized word problems when the readability level of word problems was above grade level compared to when word problems were below grade level in readability. These results suggest that personalization assists students in interpreting relatively difficult verbiage and verbal scenarios in mathematics word problems so that they can be translated into symbolic algebraic expressions (corresponding to the knowledge components labeled hard by their analysis). These results can be interpreted as partial support for both the domain interest and comprehension hypotheses.

Results of a recent observational study (blinded) of a larger sample of data from Carnegie Learning’s middle school mathematics ITS, MATHia, which is based on the Cognitive Tutor, found statistically significant, but small, associations between “honoring” preferences of students who express strong preferences (n = 518) (i.e., presenting more problems in interest areas that students rate highly, while rating other areas lowly, within a preferences dashboard) and several process variables (e.g., a variable representing the number of sections mastered per hour) found to be associated with higher standardized test scores (Fancsali and Ritter 2014). However, this study revealed that the software, in practice, was presenting problems that honored student preferences to a much smaller extent (generally only 10% to 30% of the time) than the investigators originally expected. Given this finding, one hypothesis for the lack of association between preference honoring and learning outcomes is the “personal connection” hypothesis; students do not perceive that the ITS actively takes their preferences into consideration to personalize problems.

**Conclusion**
We have reviewed several recent projects investigating the nature of improvements observed by personalizing mathematics word problem content to student interest areas outside of academics. On-going and future research will investigate further which (combination) of the domain interest, comprehension, and personal connection (and possibly other) hypotheses explain improved performance, in various contexts, for different sub-populations of students using ITSs like Cognitive Tutor. Delving into such nuance is vital to the project of determining the ways in which ITSs should adapt to learners to achieve Bloom’s “two sigma” improvement in learning. Educational technologies like ITSs are ideal for both tailoring such one-on-one personalization as well as for instrumenting, measuring, and assessing what is working in student personalization. Ideally, better understanding the driving forces behind increased performance based on non-cognitive personalization may provide basic lessons for educational psychology and practice, whether technology-enhanced or traditional.

References


