Evidence of Effectiveness

Carnegie Learning’s Cognitive Tutor® is one of the most widely researched mathematics curricula in use today. The Cognitive Tutor software has its basis in the ACT-R theory of knowledge and performance (Anderson et. al., 2004; Anderson, 2007). Following ACT-R, the Cognitive Tutor treats complex problem solving as the coordination and strengthening of a large number of relatively simple “knowledge components,” which represent strategies and concepts required to master a domain. These knowledge components are collected into a “cognitive model” which allows the software to follow individual students’ solution strategies and to track the growth of knowledge for each student over time.

The Cognitive Tutor software employs mastery learning. Students need to demonstrate mastery on each knowledge component underlying a particular topic before they can proceed to the next topic. In this way, students set their own pace through the curriculum. Mastery learning has a long history of support. Kulik et al. (1990) concluded that mastery learning programs had strong impact (average effect size of 0.5), after reviewing 108 studies.

The tutor selects problems for each student in order to maximize the amount of time that students spend on knowledge components that they have not yet mastered and to minimize the amount of time that they spend on components that they have already mastered. The tutor is able to follow individual solution strategies and to provide students with hints that are relevant to their individual approach. In addition to correct solution strategies, the cognitive model also includes information about common misconceptions and incorrect strategies and presents students with immediate feedback if they make common errors.

Researchers have conducted dozens of small-scale evaluations of individual units of instruction within the tutor. For example, Aleven and Koedinger (2002) compared two versions of the Cognitive Tutor instruction focused on reasoning about angle measures in a diagram. They found that students who were asked to articulate the reasons for geometric theorems relating angles in a diagram outscored those who were asked simply to provide the angle measures. Subsequent to these experiments, the Cognitive Tutor software was revised to include student use of geometric reasons as a fundamental part of the task. Butcher and Aleven (2008) conducted several experiments that showed that closely integrating visual reasoning about geometric diagrams with the numeric and symbolic reasoning required to determine angle measures leads to more robust learning. The most recent version of the Cognitive Tutor software incorporates these changes as well. Cen, Koedinger and Junker (2007), in a field experiment, used data from prior-year implementations to fit various parameters controlling the learning rate within the
Cognitive Tutor and found that, with properly fit parameters, students could reach the same level of performance in 12% less time.

In addition to small-scale evaluations, several larger, high-quality field trials have attested to the effectiveness of the Cognitive Tutor curriculum. The Moore Oklahoma Independent School District conducted a randomized field trial comparing instruction based on the Cognitive Tutor curriculum to control classes using the McDougal-Littell Heath Algebra I textbook (Ritter, Kulikowich, Lei, McGuire and Morgan, 2007). In three schools, teachers taught some of their classes using Cognitive Tutor and some using the textbook. Results favored the classes using Cognitive Tutor as measured by first semester grades (p=.002, d=.42), final grades (p=.007, d=.36) and scores on the ETS Algebra I end-of-course exam (p=.091, d=.38). The U.S. Department of Education’s Institute of Education Sciences’ (IES) What Works Clearinghouse recognizes this study as fully meeting their evidence standards.

Koedinger, Anderson, Hadley and Mark (1997)¹, in a matched control group quasi-experimental study of Pittsburgh Public School students taking Algebra I, found that Cognitive Tutor students outscored the control group on four separate measures: the SAT (p=.01, d=.29), the Iowa Algebra Aptitude Test (p=.02, d=.16), the problem situations test (p<.001, d=1.0) and a multiple representations test (p<.001, d=.68).

Shneyderman (2001) conducted a quasi-experimental study looking at performance of 777 students at six high schools in the Miami-Dade County Public School District. The students constituted a subset of students at those schools that had taken Algebra I and were equated at pretest both academically (using FCAT-NRT) and on various demographic characteristics (ethnicity, gender, free or reduced lunch status and limited English Proficiency). Following their Algebra I class, Cognitive Tutor students outscored the control group on the ETS Algebra I exam (p<.01, d=.22). No significant difference was observed on the end-of-year FCAT-NRT. Students in the Cognitive Tutor group passed the course at a higher rate than students in the comparison group.

Sarkis (2004) conducted a quasi-experimental study that looked at FCAT performance for 6,395 students at ten Miami-Dade County high schools, based on whether students used the Cognitive Tutor curriculum. The study found no significant difference on final grades, but it found that Cognitive Tutor students significantly outscored control group students on the FCAT (p<.001, d=.11). Results

¹ The results reported here are from an unpublished reanalysis of this data in Koedinger, Ritter and Wolfson (2007). This reanalysis used Hierarchical Linear Modeling to control for clustering and equated students at pretest for each outcome.
for students identified as having Limited English Proficiency also favored Cognitive Tutor \((p<.001, d=.28)\) as did results for students in Special Education programs \((p<.001, d=.78)\).

Plano, Ramey and Achilles (2005) reported the results of a regression discontinuity study looking at ninth grade Algebra students in the Kent, Washington School District. Students receiving below a C in their prior-year mathematics course were assigned to use Cognitive Tutor; those receiving above a C took a traditional course. The Achievement Levels Test developed by the Northwest Evaluation Association was used as both pre- and post-test to measure the improvements in learning for the two groups. The results showed that Cognitive Tutor students learned significantly more, according to this measure \((p<.05, d=.33)\). Results were reported to be even stronger for the subgroup of students identified as English Language Learners and for students receiving free or reduced lunch; since standard deviations were not included for these subgroups, effect sizes cannot be calculated.

One randomized controlled trial of the Cognitive Tutor Geometry curriculum (Pane, McCaffrey, Slaughter, Steele & Ikemoto, 2010) showed negative results \((p=.03, d=-.19)\), but the study researchers attribute the negative results to implementation problems in the urban district in which the study took place (Ikemoto, Steele, Pane & Lichter, 2012). Another consideration was the lack of focus in the Cognitive Tutor on geometric construction and proof. As a result of this finding, the software has been updated to include more focus on these topics.

Some local but high-quality studies have not found a significant advantage for Cognitive Tutor. For example, Smith (2001) reported a study of a first-year implementation of Cognitive Tutor in Virginia Beach which found no significant improvement over the control group. Similarly, Cabalo, Jaciw and Vu (2007) found no significant difference in outcomes in a study of a first-year implementation in Hawaii.

A randomized field trial of a number of educational technologies conducted by the US Department of Education (Campuzano, Dynarski, Agodini, & Rall, 2008) did not find any significant difference between students in the Cognitive Tutor condition and those in a control group in the first year of implementation. However, for the combined algebra products, there was a statistically significant increase in results when comparing teachers in their first year of implementation to these teachers in their second year of implementation, leading to an advantage for the algebra
technologies over the control group for experienced teachers (p<.05, d=.15). We believe that this finding emphasizes the need to provide teachers with training and time to adjust to the type of instruction provided by the Cognitive Tutor.

On the basis of these results, the RAND Corporation was awarded an IES Goal 4 evaluation grant, in which it evaluated the effectiveness of Cognitive Tutor Algebra 1 in one of the largest randomized control trials of its kind in seven states and 147 middle and high schools (Pane, Griffin, McCaffrey & Karam, 2014), representing a wide range of demographic characteristics. The RAND Corporation evaluation study is especially noteworthy not only for its large-scale and gold-standard design but also for its two-year approach; all but one of the previous evaluation studies have at most been conducted over one academic year, and the study conducted over two years only considered the educational effect of a pool of products in the second year. The results of the longer RAND evaluation study thus illuminates mixed results of shorter-term studies reviewed thus far.

The study compared schools who were given Cognitive Tutor texts and software to schools who continued to use the curriculum they had in place (in almost all cases, one of the top 4 Algebra textbooks). Results for high schools and middle schools were considered separately, since the middle school students were typically more advanced (taking Algebra 1 in 8th grade). Each school in the study participated in the experiment for two years, leading to two cohorts of Algebra 1 students.

In the second year of the study, both the high school and middle school Cognitive Tutor groups out-performed the control groups. The effect was only statistically significant for the high school cohort (p<.05; d=0.20) but not for the middle school cohort (p=0.16; d=0.19). Although the effect sizes were similar for both cohorts, the middle schools were smaller, resulting in less power to show a statistically significant result. The effect is large enough to be educationally meaningful. Lipsey et al. (2012) found that 8th or 9th grade students improve about 0.22 standard deviations on standardized math tests in one year. This is the improvement we expect was experienced by the control group. In the study, the Cognitive Tutor group outscored the control group by 0.20 standard deviations. This is equivalent to saying that the Cognitive Tutor group improved 0.42 standard deviations over the year, while the control group improved 0.22 standard deviations. In other words, growth in knowledge by Cognitive Tutor students nearly doubled that of the control group. The effect is also equivalent to moving a student from the 50th percentile for the Algebra I standardized test outcome to the 58th percentile.

---

2 The study was not designed to provide enough power to detect specific product effects from the first year to the second, so this result is pooled across the products in the study, rather than specific to Cognitive Tutor.
Note that these results are from the second year of the study. In the first year of the study, standardized test scores did not significantly differ between the Cognitive Tutor group and control group. In fact, the effect was slightly negative in the first year. The change between first and second year of implementation provides a context for understanding prior research on Cognitive Tutor. The student-centered, blended approach used in the Cognitive Tutor curriculum requires changes in teaching practice and support. Any major change process like this takes some time. This adjustment is not unique to Cognitive Tutor; Fullan (2002) has termed this the “implementation dip.” The RAND study shows that, although there is an adjustment period, on average, schools are able to make the adjustment in a year and, following that, achieve results superior to that which they had before. The Campuzano et al. (2008) report tells a similar story.

In contrast, Smith (2001) was a study of a first-year implementation, and the study describes how misconceptions about the intended use of Cognitive Tutor led to students transferring to other classes and other unintended effects. Cabalo et al. (2007) also report first-year adjustment problems, such as that “Teachers started using the CT textbook in mid-Fall and the CT software in January” due to difficulties delivering computers and text materials to the various islands.

Ritter et al. (2007) shows how a first-year implementation can produce positive results. This study was designed and administered by the math supervisor in the district (National Research Council, 2003), who had the ability to ensure that teachers were well trained and that the implementation was properly supported. Pane et al. (2010) demonstrates a case where the district never got to a successful implementation (Ikemoto et al., 2012).

Carnegie Learning’s long tradition of rigorous, research-based development insures that the Cognitive Tutor provides for effective and efficient student learning experiences. Collaborations with independent researchers have allowed development that is wide reaching and cutting edge. By continually incorporating new knowledge into its development, Carnegie Learning’s Cognitive Tutor is the premier intelligent tutoring system for mathematics on the market today.
References


