Cognitive Tutor: Applied research in mathematics education

STEVEN RITTER
Carnegie Learning, Inc., Pittsburgh, Pennsylvania

AND

JOHN R. ANDERSON, KENNETH R. KOEDINGER, AND ALBERT CORBETT
Carnegie Mellon University, Pittsburgh, Pennsylvania

For 25 years, we have been working to build cognitive models of mathematics, which have become a basis for middle- and high-school curricula. We discuss the theoretical background of this approach and evidence that the resulting curricula are more effective than other approaches to instruction. We also discuss how embedding a well specified theory in our instructional software allows us to dynamically evaluate the effectiveness of our instruction at a more detailed level than was previously possible. The current widespread use of the software is allowing us to test hypotheses across large numbers of students. We believe that this will lead to new approaches both to understanding mathematical cognition and to improving instruction.

For 25 years, we have been working to understand mathematical cognition through the use of cognitive modeling and applying that knowledge to constructing curricula (both text and software) that are more educationally effective than preexisting approaches. This work has been successful on many levels. It has advanced knowledge of cognition in general and of mathematical cognition in particular; the resulting curricula have proven to be educationally effective in school settings; and the curricula, as commercial products, have found a strong following in the school marketplace.

We believe that our development model, which involves a close and continuing relationship among basic research, applied research, and field testing, can serve as a model for other efforts to apply cognitive psychology to education. In this article, we describe some of the history of our efforts, our view of the relationship between basic research and development, and some directions for further research.

Background

The work that led to Carnegie Learning's Cognitive Tutors began in the psychology and computer science departments at Carnegie Mellon University. John Anderson had been developing the adaptive control of thought (ACT; later, ACT-rational, or ACT-R) theory of cognition. ACT-R is a unified theory of cognition (Newell, 1973, 1990) that aims to explain the full range of human cognition. ACT-R was implemented as a computer program, which has the advantage of requiring the theory to be precise about all of its claims. Anderson (1983) had seen great success in using

ACT-R to model laboratory results in learning, memory, and problem solving, and he was challenged to show that the same basic approach could explain cognition outside of a laboratory environment.

In its application to psychological laboratory studies, the aim of an ACT-R model is to interpret behavior. In order to interpret behavior, the model needs to correctly represent human knowledge and also to understand how that knowledge results in particular behaviors. When applied to education, this representation of knowledge results in predictions about what students can and cannot do as well as predictions about what activities and experiences will help students learn to achieve curricular goals.

The representation of knowledge inherent in this kind of model is called a *cognitive model*, and the approach of using a cognitive model in a tutoring system has come to be called a *cognitive tutor*. The first tutoring systems built in this way addressed computer programming and mathematics (Anderson, Boyle, Corbett, & Lewis, 1990; Anderson, Boyle, Farrell, & Reiser, 1987; Anderson, Conrad, & Corbett, 1989).

GPT and ANGLE, both tutors for geometry proofs (Koedinger & Anderson, 1993), were successful in a school setting. Their success, however, appeared to be highly dependent on the teacher's ability to integrate the tutoring software into broader classroom goals. This, along with Koedinger's personal experience teaching a geometry class, focused the research group on the importance of working with teachers and administrators to understand schools' curricular needs more broadly. As a consequence, the research team for the products that be-

S. Ritter, sritter@carnegielearning.com

came Carnegie Learning's Cognitive Tutors included Bill Hadley, who had taught mathematics for almost 30 years and was the 1995 recipient of the Presidential Award for Excellence in Mathematics Teaching. This team set out to build curricula that would be based on solid cognitive research, be focused on emerging national and state standards, and address the practical needs of students, teachers, and administrators.

One decision was to design a complete course, including the text, Cognitive Tutor software, ancillary materials, and training for teachers. The inclusion of text endowed the curricula with some aspects (e.g., collaboration, diagramming, and writing about mathematics) that were easier to do on paper than on the computer. The combination of text and software also helped to position the software as a regular, routine part of mathematics instruction. Instead of using the software as a "bonus" for advanced students or as a review for students who were lagging, the hybrid curricula set the expectation that software could be used as a part of primary instruction. Pilot implementations led to a model in which students used the software 2 days per week, with classroom activities structured by the text on the other 3 days each week.

The curricula proved to be educationally successful (Koedinger, Anderson, Hadley, & Mark, 1997; Koedinger, Corbett, Ritter, & Shapiro, 2000; Ritter & Anderson, 1995) and popular with students and teachers.

The Relationship Between Research and Development

The federal government's No Child Left Behind legislation mentions "scientifically based research" more than 110 times. This points to the importance of basing education on scientific research, but there are varying opinions about what it means for a curriculum to be "scientifically based." In our view, scientifically based research involves more than the demonstration that a curriculum is effective. An essential component is an explanation of *why* the curriculum is effective. Without a theoretical framework as a guide to understanding the conditions that lead to effective mathematics instruction within a curriculum, we have little hope of replicating success so that we can expand and improve instruction over time.

We think of the process of building a research-based curriculum as having four components: (1) basing the curriculum on a solid theoretical foundation, (2) applying the basic theory to the particular domain and objectives of interest, (3) evaluating results, and (4) developing and implementing a methodology for improving the curriculum on the basis of use.

Theoretical Basis

ACT-R (Anderson, 1990, 1993; Anderson et al., 2004; Anderson & Lebière, 1998) forms the primary theoretical basis of Cognitive Tutors. The primary use of the ACT-R theory has been to model important characteristics of human behavior, including error patterns and response times in studies of a variety of cognitive tasks. Most of this work has been conducted in the laboratory, but ACT-R has also been applied outside of the laboratory in areas related to

human–computer interaction, training, and education. This work has resulted in hundreds of publications (see act-r.psy .cmu.edu/publications/index.php for an extensive list).

A full explanation of ACT-R is beyond the scope of this article, but some of the tenets important to education (Anderson, 2002) include the following.

First, there are two basic types of knowledge: procedural and declarative. Declarative knowledge includes facts, images, and sounds. Procedural knowledge consists in an understanding of how to do things. All tasks involve a combination of the two types of knowledge. As we learn, we generally start out with declarative knowledge, which becomes proceduralized through practice. Procedural knowledge tends to be more fluent and automatic than declarative knowledge. Elements of procedural knowledge are referred to as *rules* or *productions* because they specify the conditions under which they are applicable and the actions (including changes in mental state) that result from applying them. Declarative knowledge tends to be more flexible and also more broadly applicable than procedural knowledge. We often refer to elements of declarative knowledge as "facts."

Second, the knowledge required to accomplish complex tasks can be described as the set of declarative and procedural knowledge components relevant to the task.

Third, both declarative and procedural knowledge become strengthened with use (and weakened with disuse). Strong knowledge can be remembered and called to attention rapidly and with some certainty. Retrieval of weak knowledge may be time-consuming, effortful, or impossible. Different knowledge components may represent different strategies or methods for accomplishing a task (including incorrect ones). The relative strength of these components helps determine which strategy is used. Learning involves the development and strengthening of correct, efficient, and appropriate knowledge components.

It is important to understand that the use of terminology in the present article differs somewhat from that in an educational context. For example, a "procedure" in ACT-R is simply a component of knowledge that can produce other knowledge components and/or lead to external behavior. In mathematics education, we might refer to the procedure of solving a linear equation. An ACT-R model of that task would consist of many productions and facts that are brought to bear. Even a simple task such as adding integers may consist of many productions, including ones associated with recalling arithmetic facts, executing counting actions, and so forth (see Lebière, 1999).

The view that emerges from ACT-R is that learning is a process of encoding, strengthening, and proceduralizing knowledge. This process happens gradually. New knowledge will be forgotten (or remain weak enough to stay unused) if it is not practiced, and elements of knowledge compete to be used on the basis of their strength (Siegler & Shipley, 1995). Since the ability to perform a task relies simply on the individual knowledge components required for that task, education is most efficient when it focuses students most directly on those individual knowledge components that have relatively low strength.

The interaction between declarative and procedural knowledge leads to an emphasis on active engagement with the conceptual underpinnings of procedures, so that students appropriately generalize this knowledge (Rittle-Johnson & Koedinger, 2002, 2005; Rittle-Johnson & Siegler, 1998; Rittle-Johnson, Siegler, & Alibali, 2001). Since procedural knowledge includes the context in which it is applicable, educational activities need to be structured so that students can practice procedures within an appropriate range of contexts.

Decomposition of complex tasks into individual knowledge components leads to a pedagogical model emphasizing practice of individual components, independent of the larger task. At the same time, some knowledge components (e.g., integration of information from smaller components) are inherent to the larger task, which provides another rationale for emphasizing performance within an appropriate context. Corbett and Anderson (1995a) report a LISP Tutor study that concludes that learning is most efficient if students master component skills first and subsequently receive scaffolding on how to integrate them into more complex tasks.

To use a sports analogy, it is important for batters to take batting practice because this will allow a baseball player to receive intensive practice with most of the skills involved in hitting a ball. However, it is also important for the batter to play in games, since some skills (e.g., reading the infield) can be practiced only in that context.

Application of Principles

Although the ACT-R theory provides a cognitive modeling framework, it does not specify the particular skills that comprise the ability to solve a linear equation, for example. In order to create instruction in mathematics, we need to understand the knowledge components involved in completing a particular task. It is not enough to know the components involved in expert performance of the task; we also need to know the components exercised by students learning to perform the task. Much of our applied research in mathematics has concerned identifying the particular skills and methods that students use to complete mathematical tasks (see Corbett, McLaughlin, Scarpinatto, & Hadley, 2000; Koedinger & Anderson, 1990; Mark & Koedinger, 1999). Often, these skills do not correspond to expert beliefs (Koedinger & Nathan, 2004; Nathan & Koedinger, 2000a, 2000b).

One technique that we have used to understand how students approach mathematics problems is to track their eye movements as they work through a problem (Gluck, 1999). Consider the task of a student completing a table of values based on a word problem such as that shown in Figure 1.

In Figure 1, part of the table that represents the word problem has been completed. The student has filled in the columns with the independent and dependent quantities relevant to the situation presented in the problem, specified the units of measurement for these quantities, and provided a formula to show their relationship. The student next needs to calculate the amount of money remaining after 2 h. There are at least two ways to perform

Figure 1. Partially completed word problem task used in an eyetracking study.

this task. First, the student might reason from the problem scenario (perhaps imagining having $20 and then using repeated subtraction to calculate the money left after spending $4 two times). A second method would be to use the algebraic expression and then substitute 2 for x and calculate the result. If a student has produced the table shown in Figure 1 (including the algebraic expression for the amount of money left), we might expect that he or she would then use the algebraic expression and execute the second method. In fact, Gluck (1999) found that when students were answering a question such as the first question in Figure 1, they looked at the problem scenario but not at the expression about 13% of the time. Students looked at the expression (sometimes along with the scenario) 54% of the time. Almost 34% of the time, they looked at neither the expression nor the problem scenario.

As a result of these and other data (see Koedinger & Anderson, 1998), the Cognitive Tutor curriculum treats the search for the algebraic expressions for simple word problems as an induction task. The formula row, shown as the second row in the table of Figure 1, is now presented at the bottom of the table, after the rows corresponding to the two questions. This has the effect of asking students to solve the individual problems (*How much money will you have after 2 hours?* and *How many hours can you play before you run out of money?*) first and then use a generalization of their reasoning to come up with the algebraic expression. In later units of curriculum, as the situations and algebraic expressions become more complex, we encourage students to go from the word problem to the expression and then to use the expression to compute specific values.

Beyond the design of mathematical tasks, the ACT-R theory guides instruction in Cognitive Tutor because the software includes an active cognitive model, which is similar to there being an ACT-R model within the software (Corbett, Koedinger, & Anderson, 1997). This model

serves two purposes. First, the model follows student actions in order to determine the particular student's strategy in solving a problem. The technique by which it does this is called *model tracing*. Second, each action that the student takes is associated with one or more skills, which are references to knowledge components in the cognitive model. Individual student performance on these skills is tracked over time (and displayed to students in the "skillometer"). Cognitive Tutor uses each student's skill profile to pick problems that emphasize the skills on which the student is weakest (Corbett & Anderson, 1995b). In addition, the skill model is used to implement mastery learning. When all skills in a section of the curriculum are determined to be sufficiently mastered, the student moves on to the next section of the curriculum, which introduces new skills.

Careful Evaluations

The development of curriculum involves many decisions, and there is often room for disagreement about how learning theory should be applied in particular cases. For that reason, we believe that careful evaluation is an essential part of the process.

Our development process has included many formative evaluations of individual units of instruction (see, e.g., Aleven & Koedinger, 2002; Corbett, Trask, Scarpinatto, & Hadley, 1998; Koedinger & Anderson, 1998; Ritter & Anderson, 1995). In addition, we have conducted several large evaluations of the entire curriculum (combining text, software, and training components in a single manipulation).

Early evaluations of Cognitive Tutors for programming and geometry showed great promise, with effect sizes of

approximately 1 *SD* (Anderson, Corbett, Koedinger, & Pelletier, 1995). In studies of the Algebra I Cognitive Tutor conducted in Pittsburgh and Milwaukee (Koedinger et al., 1997), students were tested both on standardized tests (SAT and Iowa) and on performance-based problem solving. Cognitive Tutor students significantly outscored their peers on the standardized tests (by about 0.3 *SD*s), but the difference in performance was particularly pronounced on tests of problem solving and multiple representations, on which the Cognitive Tutor students outscored their peers by 85%, representing effect sizes ranging from 0.7 to 1.2 *SD*s.

In Moore, Oklahoma, a study was conducted in which teachers were asked to teach some of their classes using Cognitive Tutor and some using the textbook they had been previously using (Morgan & Ritter, 2002; National Research Council, 2003). The result was that the Cognitive Tutor students scored higher on a standardized test (the ETS Algebra I End-of-Course Assessment), received higher grades, reported greater confidence in their mathematical abilities, and were more likely to believe that mathematics would be useful to them outside of school. This study was recognized by the U.S. Department of Education's What Works Clearinghouse as having met the highest standards of evidence. This study showed effect sizes of approximately 0.4 *SD*s.

The Miami–Dade County school district studied the use of Cognitive Tutor Algebra I in 10 high schools. An analysis of over 6,000 students taking the 2003 FCAT (a state exam) showed that students who used Cognitive Tutor significantly outscored their peers on the exam (Sarkis,
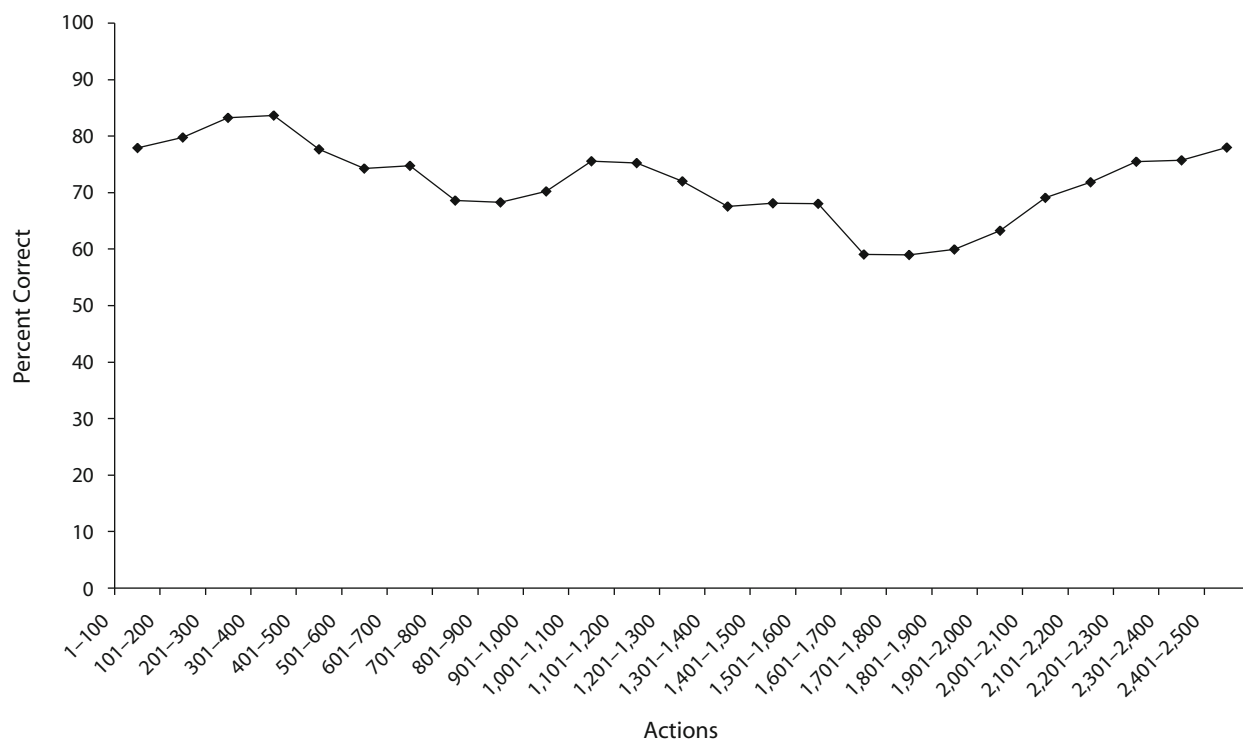

Figure 2. Percent correct over time, considering all (sequentially numbered) student actions in the geometry curriculum.
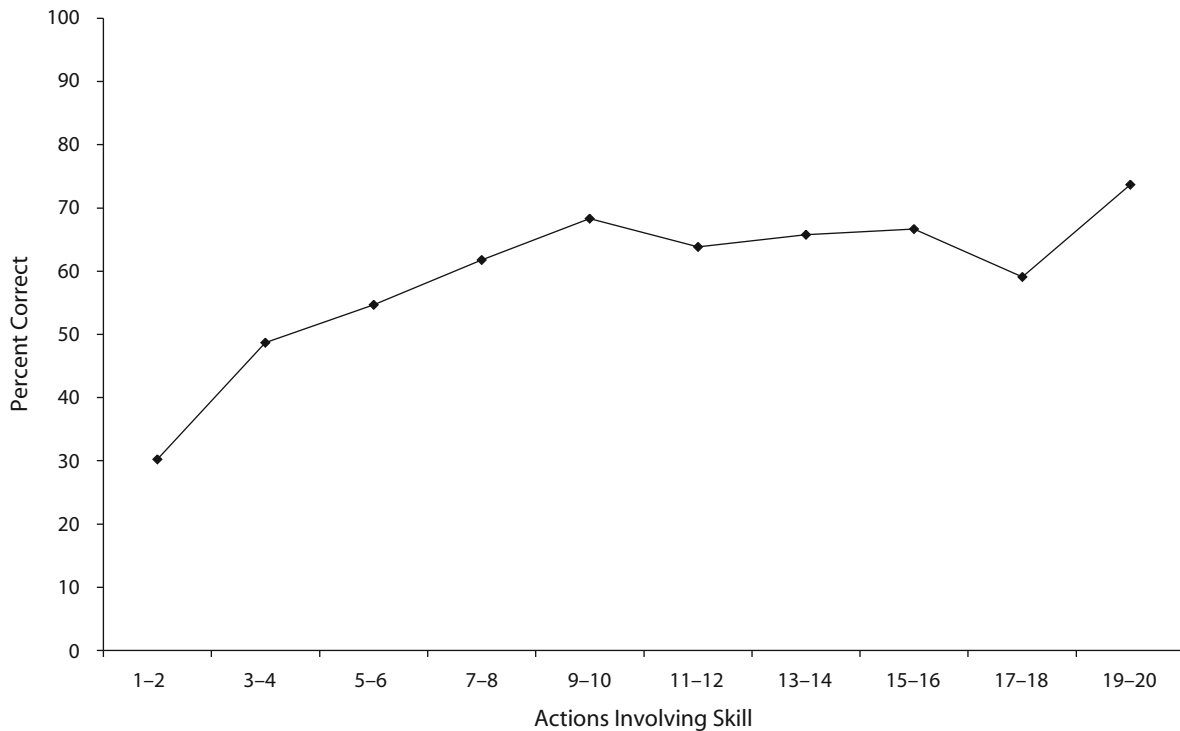
Figure 3. Percent correct over (sequentially numbered) actions involving a single skill.

2004). The findings were particularly dramatic for special populations. The study showed that 35.7% of students receiving Exceptional Student Education who use Cognitive Tutor passed the FCAT, in comparison with only 10.9% of such students who used a different curriculum. For students with limited English proficiency, 27% of Cognitive Tutor students passed the FCAT, as opposed to only 18.9% of such students in another curriculum.

Methodology for Improvement

ACT-R provides guidelines for educational pedagogy and for constructing tasks that are likely to increase learning. The theory also provides a way for us to test and improve our curriculum over time.

Cognitive Tutor observes students. As an observer, it sees everything the student does within its interface at approximately 10-sec intervals, for 2 days per week over a school year. However, the cognitive model is not a passive observer. It is continually evaluating the student and predicting what the student knows and does not know. By aggregating these predictions across students, we can test whether or not the cognitive model is correctly modeling student behavior.

Consider what an observer should see across time in a classroom. If students are learning, they should be making fewer errors over time. However, the activities given to the students over time should also increase in difficulty. In a well constructed curriculum, these two forces should cancel each other out, leading to a fairly constant error rate over time. In fact, that is what we see in the Cognitive Tutor curricula. Figure 2 shows the percent correct, over

time, for 88 students using the Cognitive Tutor Geometry curriculum in a school. The percentage correct remains fairly constant over time.

ACT-R makes the strong claim that learning takes place at the level of the knowledge components. Thus, if we consider only actions that involve a particular knowledge component, we should see an increase in percent correct over time (Anderson et al., 1989). Figure 3 shows percent correct for the same group of students as in Figure 2, this time tracking only those student actions that the cognitive model considers to be relevant to a single skill (calculating the area of a regular polygon, in an orientation in which one side is horizontal).

If ACT-R is correct in its assertion that performance of a complex task is determined by the individual knowledge components contributing to the performance of that task, then each skill in the cognitive model should show a learning curve such as this one. Failure to see learning on one of the component skills must mean that the cognitive model implemented in the tutor is not correctly representing student knowledge.

In the development of our algebra tutor, we discovered that the model was overpredicting student performance in solving some equations of the form $ax = b$. An analysis of the data revealed that the overprediction was due, in part, to the case in which $a = -1$. In retrospect, the explanation for this overprediction is obvious. In the case in which $a = -1$, the student needs to understand that the expression $-x$ means "-1 times x" and that, otherwise, the equation can be solved using the same operations as would be applied to any equation of the form $ax = b$. (An-

other way to think about this error is that some students have learned a rule equivalent to "if the equation is of the form $ax = b$, then divide by the number in front of the variable." But when the coefficient is -1, the student doesn't see a number but just a negative sign, so the rule does not apply.) Now that recognition of $-x$ as -1 times x has been added to the cognitive model, Cognitive Tutor automatically adjusts instruction to test whether or not the students have mastered that skill and automatically provides extra practice on such problems to students who need it. In addition, we can target instruction specifically to this skill.

The process of analyzing learning curves and improving our fit of these curves to the data has, to this point, been laborious. We have recently been exploring the possibility of automating the process of discovering flaws in the cognitive model (Cen, Koedinger, & Junker, 2005; Junker, Koedinger, & Trottini, 2000), and this is an active focus of research at the Pittsburgh Science of Learning Center (www.learnlab.org).

We believe that in the near future we will be able to greatly extend our ability to understand and accurately model students' mathematical cognition. In addition to improved statistical modeling techniques, the expansion of Carnegie Learning's customer base and the ability to aggregate student data over the Internet provides us with the ability to look at student cognition both more deeply and more broadly.

We have now collected data from over 7,000 students using Cognitive Tutor in a pre-algebra class. These data comprise over 35 million observations, which amounts to observing an action for each student about every 9.5 sec. With a database of this size, we expect to be able to detect subtler factors affecting learning, including the effectiveness of individual tasks, hints, and feedback patterns. We are starting to apply microgenetic methods (Siegler & Crowley, 1991) to see whether or not we can identify key learning experiences, which could contribute to better cognitive models of individual differences in prior knowledge or learning styles and preferences. We believe that the combination of a dense data stream of student behavior and a large sample of students will allow us to greatly expand our knowledge of students' mathematical cognition and advance our ability to help students learn mathematics.

AUTHOR NOTE

Correspondence concerning this article should be addressed to S. Ritter, Carnegie Learning, Frick Building, 20th Floor, 437 Grant Street, Pittsburgh, PA 15219 (e-mail: sritter@carnegielearning.com).

REFERENCES

Aleven, V. A. W. M. M., & Koedinger, K. R. (2002). An effective metacognitive strategy: Learning by doing and explaining with a computer-based Cognitive Tutor. *Cognitive Science*, **26**, 147-179.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.

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

Anderson, J. R. (2002). Spanning seven orders of magnitude: A challenge for cognitive modeling. *Cognitive Science*, **26**, 85-112.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebière, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, **111**, 1036-1060.

Anderson, J. R., Boyle, C. F., Corbett, A. T., & Lewis, M. W. (1990). Cognitive modeling and intelligent tutoring. *Artificial Intelligence*, **42**, 7-49.

Anderson, J. R., Boyle, C. F., Farrell, R., & Reiser, B. J. (1987). Cognitive principles in the design of computer tutors. In P. Morris (Ed.), *Modelling cognition* (pp. 93-133). Chichester, U.K.: Wiley.

Anderson, J. R., Conrad, F. G., & Corbett, A. T. (1989). Skill acquisition and the LISP tutor. *Cognitive Science*, **13**, 467-505.

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

Anderson, J. R., & Lebière, C. (1998). *The atomic components of thought*. Mahwah, NJ: Erlbaum.

Cen, H., Koedinger, K. R., & Junker, B. (2005). Learning Factors Analysis: A general method for cognitive model evaluation and improvement. In M. Ikeda, K. Ashley, & T. Chan (Eds.), *Intelligent Tutoring Systems 8th International Conference* (pp. 164-175). Berlin: Springer.

Corbett, A. T., & Anderson, J. R. (1995a). Knowledge decomposition and subgoal reification in the ACT programming tutor. In J. Greer (Ed.), *Artificial intelligence and education, 1995: The proceedings of AI-ED 95* (pp. 469-476). Charlottesville, VA: AACE Press.

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

Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Intelligent tutoring systems. In M. G. Helander, T. K. Landauer, & P. Prabhu (Eds.), *Handbook of human–computer interaction* (2nd ed., pp. 849-874). Amsterdam: Elsevier.

Corbett, A. [T.], McLaughlin, M., Scarpinatto, K. C., & Hadley, W. H. (2000). Analyzing and generating mathematical models: An Algebra II cognitive tutor design study. In G. Gauthier, C. Frasson, & K. van Lehn (Eds.), *Intelligent Tutoring Systems: Fifth international conference* (pp. 314-323). Berlin: Springer.

Corbett, A. T., Trask, H. J., Scarpinatto, K. C., & Hadley, W. S. (1998). A formative evaluation of the PACT Algebra II Tutor: Support for simple hierarchical reasoning. In B. P. Goettl, H. Halff, C. Redfield, & V. Shute (Eds.), *Intelligent Tutoring Systems: Fourth International Conference, ITS '98* (pp. 374-383). New York: Springer.

Gluck, K. A. (1999). Eye movements and algebra tutoring. *Dissertation Abstracts International*, **61**, 1664B.

Junker, B. W., Koedinger, K. R., & Trottini, M. (2000, July). *Finding improvements in student models for intelligent tutoring systems via variable selection for a linear logistic test model*. Paper presented at the 65th Annual Meeting of the Psychometric Society, Vancouver.

Koedinger, K. R., & Anderson, J. R. (1990). Abstract planning and perceptual chunks: Elements of expertise in geometry. *Cognitive Science*, **14**, 511-550.

Koedinger, K. R., & Anderson, J. R. (1993). Effective use of intelligent software in high school math classrooms. In *Proceedings of the Sixth World Conference on Artificial Intelligence in Education* (pp. 241-248). Charlottesville, VA: Association for the Advancement of Computing in Education.

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., Anderson, J. R., Hadley, W. H., & Mark, M. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education*, **8**, 30-43.

Koedinger, K. R., Corbett, A. T., Ritter, S., & Shapiro, L. J. (2000). *Carnegie Learning's Cognitive Tutor: Summary research results*. Pittsburgh: Carnegie Learning. Available at www.carnegielearning.com/web_docs/CMU_research_results.pdf.

Koedinger, K. R., & Nathan, M. J. (2004). The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences*, **13**, 129-164.

Lebière, C. (1999). The dynamics of cognition: An ACT-R model of cognitive arithmetic. *Kognitionswissenschaft*, **8**, 5-19.

Mark, M. A., & Koedinger, K. R. (1999). Strategic support of algebraic expression writing. In *Proceedings of the Seventh International*

Conference on User Modeling (pp. 149-158). Available at www .cs.usask.ca/UM99/Proc/mark.pdf.

Morgan, P., & Ritter, S. (2002). *An experimental study of the effects of Cognitive Tutor Algebra I on student knowledge and attitude*. Pittsburgh: Carnegie Learning. Available at www.carnegielearning .com/web_docs/morgan_ritter_2002.pdf.

Nathan, M. J., & Koedinger, K. R. (2000a). An investigation of teachers' beliefs of students' algebra development. *Cognition & Instruction*, **18**, 209-237.

Nathan, M. J., & Koedinger, K. R. (2000b). Teachers' and researchers' beliefs about the development of algebraic reasoning. *Journal for Research in Mathematics Education*, **31**, 168-190.

National Research Council (2003). *Strategic education research partnership* (M. S. Donovan, A. K. Wigdor, & C. E. Snow, Eds.). Washington, DC: National Academies Press.

Newell, A. (1973). You can't play 20 questions with nature and win: Projective comments on the papers of this symposium. In W. G. Chase (Ed.), *Visual information processing* (pp. 283-310). New York: Academic Press.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Ritter, S., & Anderson, J. R. (1995). Calculation and strategy in the equation solving tutor. In J. D. Moore & J. F. Lehman (Eds.), *Proceedings of the 17th Annual Conference of the Cognitive Science Society* (pp. 413-418). Hillsdale, NJ: Erlbaum.

Rittle-Johnson, B., & Koedinger, K. R. (2002). Comparing instructional strategies for integrating conceptual and procedural knowledge. In D. S. Mewborn, P. Sztajin, D. Y. White, H. G. Wiegel, R. L. Bry-

ant, & K. Nooney (Eds.), *Proceedings of the 24th Annual Meeting of the North American Chapters of the International Group for the Psychology of Mathematics Education* (pp. 969-978). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Rittle-Johnson, B., & Koedinger, K. R. (2005). Designing better learning environments: Knowledge scaffolding supports mathematical problem solving. *Cognition & Instruction*, **23**, 313-349.

Rittle-Johnson, B., & Siegler, R. S. (1998). The relation between conceptual and procedural knowledge in learning mathematics: A review. In C. Donlan (Ed.), *The development of mathematical skills: Studies in developmental psychology* (pp. 75-110). Hove, U.K.: Psychology Press.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of Educational Psychology*, **93**, 346-362.

Sarkis, H. (2004). *Cognitive Tutor Algebra 1 program evaluation: Miami–Dade County public schools*. Lighthouse Point, FL: The Reliability Group. Available at www.carnegielearning.com/web_docs/sarkis_2004 .pdf.

Siegler, R. S., & Crowley, K. (1991). The microgenetic method: A direct means for studying cognitive development. *American Psychologist*, **46**, 606-620.

Siegler, R. S., & Shipley, C. (1995). Variation, selection, and cognitive change. In T. J. Simon & G. S. Halford (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 31-76). Hillsdale, NJ: Erlbaum.