How Handwritten Input Helps Students Learning Algebra Equation Solving

Lisa Anthony **Jie Yang** **Kenneth R. Koedinger**

March 1, 2008
CMU-HCII-08-100

School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

This work is supported by the NSF Award #SBE-03554420 and the Pittsburgh Science of Learning Center.

Keywords: Math learning, handwriting input, intelligent tutoring systems

Abstract

Building on past results establishing a benefit for using handwriting when entering mathematics on the computer, we hypothesize that handwriting as an input modality may be able to provide significant advantages over typing in the mathematics learning domain. The use of handwriting may result in decreased extraneous cognitive load on students, and it may provide better support for the two-dimensional spatial components of mathematics when compared to existing typing-based tools. It may also simply make for an easier transition to paper and speed up entry of mathematics, and thus free up more student time for learning. Here we report the results of a study in which middle and high school students used a software tutor for algebra equation solving with either typing or handwriting as the input modality. We found that the handwriting modality resulted in similar learning gains in much less time than the typing modality. We also found that students seem to experience a higher degree of transfer in the handwriting modality than in the typing modality based on performance during training. An implication of these results is that students could achieve farther goals in an intelligent tutoring system curriculum when they use handwriting interfaces *vs.* typing. Both of these results represent promising findings and encourage future exploration of the use of handwriting interfaces for mathematic instruction online.

1 Introduction

Many schools throughout the United States now incorporate computers as a regular part of classroom instruction [12] and use intelligent tutoring systems as supplements to traditional classroom instruction. An intelligent tutoring system is educational software that can monitor the student as he/she works at his/her own pace, and tailor feedback, step-by-step hints, and even the curriculum to address the student's particular needs. This self-pacing provides an opportunity for teachers to give more individual attention to students that need it most.

Although intelligent tutors for math have improved with respect to pedagogical style and overall effectiveness over the last 15 years (*e.g.*, [3]), their interfaces have remained more or less the same: keyboard-and-mouse windows-icons-menus-pointing (WIMP) interfaces. *Output* modality contrasts have been studied with respect to learning, including the use of animations, diagrams and talking heads (*e.g.*, [5], [7], but the literature has been silent on the effects of *input* modality on learning[1]. We believe that the input modality is extraneous to the problem-solving process. WIMP interfaces may impose extraneous cognitive load on the student, because representing and manipulating mathematics equations can be cumbersome in a typing interface. An interface that can more directly support the standard notations for the mathematics the student is learning would reduce extraneous cognitive load and lead to increased learning (*c.f.*, [10]).

This paper reports evidence in favor of handwriting-based interfaces with respect to learning in the domain of algebra equation solving. The study reported specifically addresses the speed and user satisfaction benefits of handwriting found for college-age students in [2] as well as learning-specific measures. The results show that handwriting input continues to have benefits when extended to middle and high school students engaged in a learning task. We propose that deeper explorations are needed in order to establish a theoretic foundation on how to achieve better learning gains using an appropriate interface.

2 Background and Motivation

Intelligent tutoring systems are beginning to explore more natural interfaces such as natural language processing of typed input (*e.g.*, [1]) and spoken dialogues with conversational agents (*e.g.*, [5]). Most systems still currently rely on standard WIMP interfaces, though. This is due in part to the fact that the technology available to most students in the classroom is limited to keyboard-and-mouse-this situation is changing however, as students receive PDAs or TabletPCs in the classroom [12]. However, while advantages of pen-based input have been explored for the math domain in terms of usability measures such as speed and user satisfaction [2], very little work has been done analyzing the effect of modality on learning. One study has reported results comparing a variety of pen-based interfaces for solving geometry problems with students [8], but it assumes that handwriting is beneficial and does not provide a current practice (typing) control condition for comparison.

[1]Note that input modality here refers to the modality of generation by the student, and the output modality is the modality presented to the student by the system.

There is evidence that the use of handwriting interfaces could have particular pedagogical advantages in the domain of learning environments, especially for the mathematics domain. Prior work on entering equations using different modalities including typing, handwriting, and speech has indicated that handwriting is the faster and favored modality to typing [2]. The increased efficiency of a handwriting interface for a math tutor would allow students to accomplish more problems in the same amount of time, and the fact that students prefer handwriting might lead to increased engagement during tutoring (*c.f.*, [4]). Another factor is that in mathematics, the spatial relationships among symbols have inherent meaning. For example, the placement of the x in the following two expressions significantly changes the meaning of the expression: $2x$ *vs.* 2^x. Handwriting is a much more flexible and robust modality for representing and manipulating such spatial relationships, which become more prevalent as students advance in math training to calculus and beyond. Further, students practice in the classroom and on homework and take tests on paper using handwriting; this modality becomes more *fluent* for students when solving algebra equations. An interface which can take advantage of this should allow a higher degree of transfer and cause the tutoring system to overpredict student performance after achieving mastery in a lesson less than a typing interface for the same lesson.

3 Experimental Method

Previous studies have looked at the effect of input modality on usability [2]; we were interested in determining whether or not similar effects would occur in a learning domain. Do students experience differences in learning due to the modality in which they generate their answers? In addition, we wanted to see whether the effects reported in [2] for college students entering calculus expressions with complex symbols often not found on keyboards, would generalize to a younger population and simpler equations that can be typed easily.

We explored three modalities: *typing*, in which students typed out the solution in a blank text box; *handwriting*, in which students wrote the solution using a stylus in a blank space on the screen; and *handwriting-plus-speaking*, identical to handwriting but students were also asked to speak aloud the steps to the problems they were solving. We included this condition based on prior work finding that spoken self-explanations are more effective for learning than written or typed ones [6].

3.1 Participants

In this study, 48 middle and high school students participated. Ten students had to be dismissed due to technical difficulties with the experiment software or due to scoring 100% on the pre-test measuring algebra skills. Of the remaining 38 students, 19 participants were female and 19 were male. They ranged from 6th to 10th grades (ages 11 to 17, mean age=13.5 yrs). They were all paid participants who responded to a newspaper ad offering help in algebra. Most students had not used handwriting input on the computer before and two-thirds claimed to be very comfortable with typing. In spite of the wide range of ages and grades, most students were at about the same level

$$\frac{28-2y}{9-(4y+c)} = 6y+4$$

Copying Phase

$$\frac{x}{5}+4=7$$

Learning Phase

$$\frac{56}{x}=7$$
$$x\cdot\frac{56}{x}=7\cdot x$$
$$56=7x$$
$$\frac{56}{7}=\frac{7x}{7}$$
$$8=x$$

Sample Worked Example from Learning Phase

Figure 1: Samples of equations and problems from each phase of the experiment.

of algebra skills. No effects of ethnicity, gender, age or other demographic data were seen during exploratory data analysis.

3.2 Procedure

All students came to a research laboratory at the university for about 2.5 hours. The session had two main phases. In the first phase, students copied given equations of beginning algebra level (Figure 1). Each student copied equations in all three conditions. In the second part, following a brief pre-test to gauge prior algebra knowledge, students solved beginning algebra equations in one of the three conditions. The equations students saw in this part of the session were simpler than in the copying phase (Figure 1). During the problem-solving phase, students alternated copying a non-annotated worked example (Figure 1) and then solving an analogous equation while referring to the example. This instructional paradigm was modeled after [11] and was chosen because we did not provide step-by-step feedback *during* problem-solving due to technological constraints of recognizing handwritten input. The example was intended to provide a kind of step-by-step *feed-forward* to aid students. When students completed a problem their answer was sent to an experimenter at a separate computer for answer verification; the experimenter responded only Yes or No based on the student's final answer and this response was then shown on the student's screen. Students were not given specific reasons why their answer was incorrect. After 3 incorrect attempts, the program automatically displayed the correct solution; the students copied it and moved on. We controlled for content rather than time in this study. When the students had completed all 9 problems in the curriculum, they took a paper post-test on problems that were isomorphic to the study problems. They then filled out a questionnaire about their satisfaction with their experiences in the session, their prior math and computer skills, and what their favorite modality was in the study.

3.3 Measures

Dependent variables in this study were different in each phase of the session. In the copying phase, they included the time it took students to copy each equation and the number of errors they made during copying. In the solving phase, the dependent variables included the total time it took students to complete all problems; the time it took them to solve each problem or copy each example; the number of attempts it took them during training to either get the answer correct or move on (max=3); and the change in score from pre-test to post-test. General dependent variables included the responses on the user satisfaction questionnaire about the three modalities they had tried. In this paper we focus on time, test scores, attempts during training and satisfaction.

4 Results and Discussion

4.1 Time on Task

We measured the time students took to complete both the copying phase and the learning phase. During both the copying and the learning phases, students took about twice as long to enter equations in typing than in the other two conditions. In a repeated measures analysis of the within-subjects factor of condition with the dependent variable of average time per equation in the copying phase, we found a significant main effect of the within-subjects factor of condition ($F_{2,74} = 49.60$, $p < 0.0005$), with a planned contrast showing typing as the slowest condition ($F_{1,37} = 58.49$, $p < 0.0005$). In a univariate ANOVA on total time to solve the problems with condition as a fixed factor in the learning phase, we found a significant main effect of condition ($F_{2,35} = 11.05$, $p < 0.0005$) in which typing was the slowest ($t(35) = 4.70$, $p < 0.0005$). This supports the findings in [2] that showed handwriting was the faster modality when compared to typing; the effect is smaller in this case (two times faster here *vs.* three times faster in the earlier study), which may be due to the complexity of the interface used in the prior study (Microsoft Equation Editor). The time-speedup during the learning phase is not a direct measure of learning gain, but has implications for learning in that students who can get through problems more quickly by virtue of a more natural interface can therefore advance farther in the curriculum than if they had been typing.

We hypothesized that equations with 2D elements such as fractions would impact the time taken and/or learning experienced for the typing modality and not the handwriting modality. We found in both the learning and the copying phases that this was true. In the copying phase, about 40% of the equations contained fractions. A separate repeated measures analysis on time per equation (on the 26 students for whom we had data on what problems contained fractions in the copying phase) revealed a significant interaction between the two within-subjects factors of condition and appearance of fractions ($F_{1,25} = 4.76$, $p < 0.05$). In the learning phase, half of the problems contained fractions and half did not. A repeated measures analysis of the average time students took per problem to solve problems with fractions *vs.* without fractions revealed a significant interaction between the between-subjects factor of condition and the within-subjects factor of appearance of fractions ($F_{2,36} = 5.252$, $p < 0.01$). Figure 2 shows the interaction plot of appearance of fractions and input condition for the learning phase. The typing condition is slowed down more by the

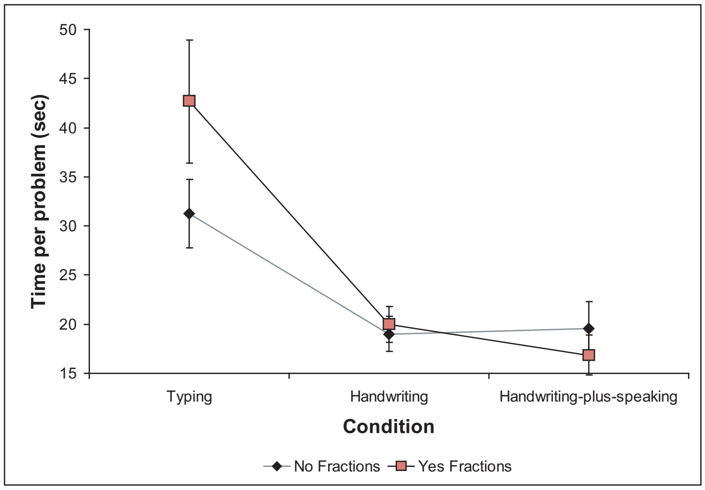

Figure 2: Average time per problem by condition crossed with appearance of fractions in the learning phase for both copying examples and solving problems. Error bars show 95% confidence interval.

appearance of fractions, whereas there is no difference in the other two conditions when fractions appear *vs.* when they do not. This interaction between speed and complexity of the math replicates the findings in [2] that showed a similar interaction based on appearance of non-keyboard characters. The result reported in this study is more robust in that the interaction is based on spatial characteristics of the math to be input rather than simply what is easily typeable on a keyboard. This result implies that the speed benefits of handwriting input will magnify as students progress to more complex math such as polynomial algebra or calculus, which contain high frequencies of fractions, exponents, etc. Again, students may be able to progress farther more quickly by virtue of handwriting's faster, more natural support for these notations.

4.2 User Preferences

All students were exposed to all three conditions during the copying phase. Students showed a strong preference for handwriting. Out of 38 total students, only 21% said keyboard/typing was their favorite method, while over 78% preferred one of the methods with handwriting. As shown in Figure 3, this difference was not based on a bias to prefer the method used during learning (a Chi-Square test of independence reveals no significant association; Pearson coefficient$= 1.802$, $p = 0.77$). This lack of association implies little or no novelty effect of handwriting was seen. A variety of typical qualitative comments from some of the students are included in Table 1.

4.3 Learning Gains and Learning Efficiency

Despite taking about half the time during the learning phase, the handwriting students learned just as much as the typing students. There was no significant difference among the conditions with

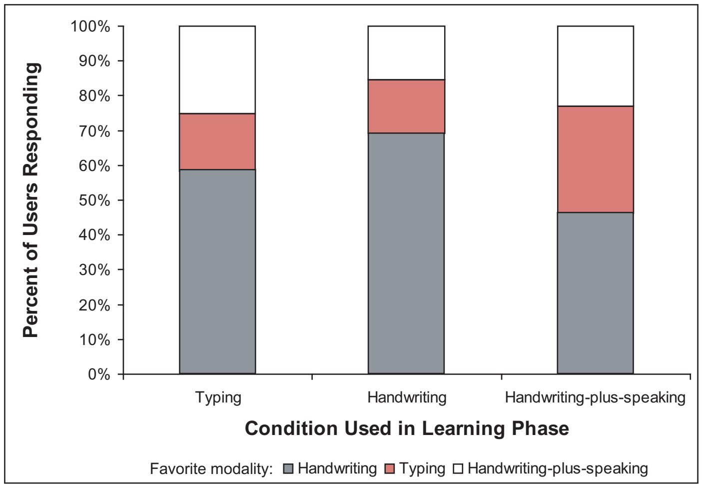

Figure 3: Histogram of user responses rating their *favorite* input modality grouped by the modality they used during the learning phase of the session.

Table 1: Comments made by students on the post-session questionnaire about each modality.

Typing	**Handwriting**	**Handwriting-plus-speaking**
"It took too long and was hard to get everything where I wanted."	"Yes, because it is how I'm used to doing problems in math class, by writing them out."	"It made it easier to think it out when I said it while doing it."
"It takes me longer to type math problems [as opposed] to [writing] them."	"It is easier than typing."	"[It's] easier to understand when you talk through the problems."
	"It was better than typing."	"I like talking through the problems it made me focus more."
	"It was a lot easier and I finished quickly."	

respect to the learning gain from pre-test to post-test ($F_{2,35} = 0.293$, *n.s.*). This means that, even though students solved the same amount of problems and took less time in handwriting than in typing, their learning as measured by performance improved about the same amount ($mean = 11.75\%$, $stdev = 17.34$). This measure of learning is relatively coarse; in future studies we intend to analyze in more detail the concepts students mastered rather than purely raw gain scores which do not reflect *how* the learning may have differed among conditions.

Although learning gains appeared to be of the same magnitude based on pre- to post-test scores, the fact that the time spent per condition was so different suggests that perhaps handwriting was a more *efficient* learning modality than typing. The concept of *learning efficiency* has been used in, for example, [9], to explore how students may be able to achieve similar levels of mastery but do fewer problems. In this study, students all did the same number of problems, but the time spent per session was significantly different by modality. If time in each modality is significantly different, it follows that learning *rate per hour* should be significantly different. We extrapolated the pre- to post-test gain per hour of instruction for each modality based on the average time over all three conditions per session for the learning phase (31 minutes). Handwriting performed the best with 110% gain per hour, while typing and handwriting-plus-speaking are about equal with 63% and 72% gain per hour, respectively. While there is a trend in favor of handwriting, we were unable to determine a method of measuring learning efficiency that accurately accounted for difficulties in the measurement of learning gain in this study, such as the fact that some students had negative gain scores. We believe it is safe to assume that no student actually *lost* knowledge during our experiment; therefore, negative gain scores are most likely due to noise in the measurement. It is unclear how other approaches to measuring learning efficiency account for such difficulties. We will continue to explore this issue in further studies.

4.4 Transfer to Paper

One of our hypothesized advantages of using the handwriting modality is that handwriting will allow a greater degree of transfer to paper than using typing interfaces. In this study we attempted to assess level of transfer in each condition by correlating the pre-test score and post-test score with performance during training. We hypothesized that the cases in which there was a *modality switch* (*i.e.*, writing on the pre-test to typing in the interface to writing on the post-test) should have a lower correlation in performance during training *vs.* on the tests. We ran bivariate correlations of percent of problems solved on the first try during training and the *pre-test* score, grouped by condition. The Pearson correlation for the typing condition was not statistically significant (0.343, $p = 0.275$), whereas for the two handwriting conditions, there was a significant correlation (0.613, $p < 0.05$ for handwriting; 0.614, $p < 0.05$ for handwriting-plus-speaking). We then ran separate bivariate correlations of training performance and the *post-test* score. The Pearson correlation for the typing condition in this case was also not statistically significant (0.320, $p = 0.310$), whereas for the two handwriting conditions, there was a significant correlation (0.708, $p < 0.01$ for handwriting; 0.553, $p = 0.05$ for handwriting-plus-speaking). These results show that handwriting does indeed afford students a higher degree of transfer to paper because it does not involve a modality switch from training to testing. Performance during testing more closely matches performance during training when the modality of testing is similar to that of training (or vice versa).

4.5 Handwriting-plus-Speaking

Throughout this paper, we have focused primarily on contrasting the typing condition with both handwriting conditions. However, the handwriting-plus-speaking condition illustrated a trade-off between speed and learning that did not manifest in the other two conditions. While the multimodal students were about as fast as the handwriting students, they did not learn as much, based on the time-adjusted score. (They did learn more than the typing students.) It appears from these preliminary data that the added cognitive load of speaking somehow interferes with learning the goal concept. Teachers frequently speak and write math while lecturing at the blackboard, yet it seems that students are not comfortable with generating solutions in both modalities at once. Prior literature has demonstrated that students learn better when they self-explain, and even further, that they learn better when their self-explanations are spoken rather than typed [6]. Our results indicate that the learning benefits for students self-explaining aloud do not manifest in a "shadowing" paradigm in which students simply state aloud the problem-solving steps they perform. This can inform future design of instructional paradigms that support or require student self-explanations.

5 Conclusions and Future Work

In this paper we have reported a study that provides valuable early evidence in favor of handwriting-based interfaces for intelligent tutoring systems in mathematics, especially algebra equation solving. Students are able to solve problems twice as quickly in handwriting than in typing by virtue of increased input speed. Despite this much faster progress through the lesson, students seem to learn just as much as their typing counterparts based on test performance. This implies that the increased time spent typing is not valuable to the learning process. In addition, students seem to achieve a higher degree of transfer when using handwriting on the computer than when using typing. Finally, most students choose handwriting as their favorite input modality, citing its "ease" and "naturalness" and similarity to the paper notations with which they are already familiar.

This study is an early laboratory study in a program of research to explore the factors that contribute to handwriting's advantages for learning in this domain. It has several limitations which will be addressed in future studies in real-world classrooms. Specifically, the typing interface provided was simple and not representative of existing intelligent tutoring systems. We are currently developing a study that will compare handwriting interfaces with the current state-of-the-art by comparing our prototype system with existing Cognitive Tutors (*e.g.*, [1]). We also plan to focus more directly on issues of cognitive load and isolating the effects of the modality on cognitive load which are extraneous to learning. Although student performance from pre- to post-test did not differ significantly between conditions, other learning measures may reveal differences in what and how the students learned. For instance, analyzing the number of errors during training may reveal ways in which typing may impose more extraneous cognitive load than handwriting. Finally, future studies will control for mastery as in the real classroom, rather than concepts covered, so we can compare progress and learning between modalities given equal time more directly.

6 Acknowledgments

The authors wish to thank Tom Bolster, Noboru Matsuda, Chris Atkeson, Shelley Evenson, and Sue Fussell for their invaluable equipment, time, and support.

References

[1] V.A.W.M.M. Aleven, K.R. Koedinger, and O.A. Popescu. A tutorial dialog system to support self-explanation: Evaluation and open questions. In *Proceedings of the International Conference on Artificial Intelligence in Education*, pages 39–46, 2003.

[2] L. Anthony, J. Yang, and K.R. Koedinger. Evaluation of multimodal input for entering mathematical equations on the computer. In *Proceedings of the ACM Conference on Human Factors in Computing (CHI)*, pages 1184–1187, 2005.

[3] A. Corbett and H. Trask. Instructional interventions in computer-based tutoring: Differential impact on learning time and accuracy. In *Proceedings of the ACM Conference on Human Factors in Computing (CHI)*, pages 97–104, 2000.

[4] C.S. Dweck and E.L. Leggett. A social-cognitive approach to motivation and personality. *Psychological Review*, 95:256–271, 1988.

[5] A. Graesser, K.N. Moreno, J.C. Marineau, A.B. Adcock, A.M. Olney, and N.K. Person. Autotutor improves deep learning of computer literacy: Is it the dialog or the talking head? In *Proceedings of the International Conference on Artificial Intelligence in Education*, pages 47–54, 2003.

[6] R.G.M Hausmann and M.T.H. Chi. Can a computer interface support self-explaining? *Cognitive Technology*, 7:4–14, 2002.

[7] R.E. Mayer. *Multimedia Learning*. Cambridge University Press, Boston, MA, 2001.

[8] S. Oviatt, A. Arthur, and J. Cohen. Quiet interfaces that help students think. In *Proceedings of the ACM Symposium on User Interface Software and Technology*, pages 191–200, 2006.

[9] M. Ringenberg and K. VanLehn. Scaffolding problem solving with annotated worked-out examples to promote deep learning. In *Proceedings of the International Conference on Intelligent Tutoring Systems*, pages 625–634, 2006.

[10] J. Sweller. Cognitive load during problem solving: Effects on learning. *Cognitive Science*, 12:257–285, 1988.

[11] J.G. Trafton and B.J. Reiser. The contributions of studying examples and solving problems to skill acquisition. In *Proceedings of the Cognitive Science Society*, pages 1017–1022, 1993.

[12] C. Wood. Technology and education. PC Magazine: http://www.pcmag.com/article2/0,4149,15154,00.asp, 2002.