A Cognitive Analysis of Equation Reading Applied to the Development of Assistive Technology for Visually-Impaired Students

Article *in* Proceedings of the Human Factors and Ergonomics Society Annual Meeting · September 2004

DOI: 10.1177/154193120404800538

Douglas Gillan
North Carolina State University

Arthur I. Karshmer
University of San Francisco

A COGNITIVE ANALYSIS OF EQUATION READING APPLIED TO THE DEVELOPMENT OF ASSISTIVE TECHNOLOGY FOR VISUALLY-IMPAIRED STUDENTS

Paula Barraza[1], Douglas J. Gillan[1], Arthur Karshmer[2], and Skye Pazuchanics[1]
New Mexico State University[1]
University of South Florida[2]

The purpose of this research was to investigate the perceptual and cognitive processes involved in equation reading to apply that knowledge to the development of assistive technology for blind equation readers. The research used a process tracing observational study, three experiments, and an eye-tracking study to examine several hypotheses about equation reading: people (1) read equations from left to right, one element at a time, (2) back scan when reading equations, (3) substitute the outcome of a parenthetical expression for the initial elements, and (4) scan the entire equation before element by element reading to create a schematic structure. The process tracing study provided evidence for all of the hypotheses, with the experiments supporting the first three hypotheses, but not the fourth. These results have been implemented in assistive software for visually-impaired users, the Math Genie -- an auditory browser.

INTRODUCTION

Mathematical relations can be represented visually in various formats -- graphs, tables, and equations. Equations present the relation by use of symbols in a highly structured organization. Skill in math is based, in part, on the ability to read and solve equations accurately. Because of the importance of visual structure for equation reading, presenting the same information auditorially would likely increase the difficulty in reading and solving an equation. Therefore, conditions in which people receive equations only via the auditory channel might put the hearer at a disadvantage. The conditions that would involve auditory presentation of equations include both equation reading for persons who are blind and those with low vision, but also fully sighted classmates reading math homework over the phone. And a strictly linear auditory presentation of the equation would particularly disadvantage the hearer. Unfortunately, most auditory browsers (and probably most classmates reading over the phone) present equations strictly linearly. Fortunately, an auditory browser need not be restricted to a strict linear presentation of the symbols in an equation.

Our approach to browser design has been (1) to investigate the perceptual and cognitive processes used in equation reading by sighted people, then (2) to provide the same process capabilities in the auditory channel to blind readers that sighted readers have via the visual channel. Thus, to the extent that sighted readers use approaches that depart from linear reading of the equation, the auditory browsers will have the capability to provide hearers with nonlinear approaches to equation reading.

How might people read equations? One hypothesis is that sighted people read equations in the same way that they read text. People read text primarily letter-by-letter beginning at or near the leftmost symbol in a text line and proceeding to the right; readers also often engage in backward scans to retrieve information that was previously read (Smyth, Collins, Morris, & Levy 1994). Alternatively, one might propose that, given the different goals for reading equations and reading text, users interact with equations and texts in very different ways. Specifically, they may engage in an initial scan of the equation to gain information that would allow the reader to set up a schematic representation of the equation structure, followed by an element-by-element reading that fills in the slots in the schema.

STUDY: PROCESS TRACING

We began by observing 15 students as they read equations; they were instructed to think aloud as they read the equations and received practice with the think aloud protocol. The 14 equations that they read varied in the types and number of arithmetic operators and in the overall complexity of the equation. The verbal protocols of the participants' thinking aloud were analyzed to identify the cognitive and perceptual processes as they read and solved equations. The analysis suggested the following features of equation reading: (1) reading proceeds largely in a left-to-right direction reading one element at a time, much like reading text; (2) on approximately half of the observed equations, readers performed an initial scan to determine the structure of the problem; (3) readers frequently engaged in backward scanning to previously-read elements; (4) readers chunked together elements of the equation (usually within parentheses) and solved those chunks as separate modules; and (5) as a consequence of chunking, readers solved the equation hierarchically.

Because verbal protocol data may, under some conditions, reflect the participants' expectations concerning their cognitive processes, rather than the actual processes (e.g., Nisbett and Wilson, 1977), conclusions drawn from analysis of verbal protocols should be tested by means of true experiments. Accordingly, we conducted a series of experiments to further examine the five features of equation reading described above

EXPERIMENT 1: EQUATION RECALL

Method

In Experiment 1, 11 participants read 72 computer-displayed equations varying in complexity (24 low, 24 medium, and 24 high), one at a time. On each trial, the participant first read the equation (for 1, 2, or 4 sec), followed by a distracter screen consisting of dots (with a duration of 1 sec or 8 sec). Then, participants were asked to recall the equation by writing down everything that they could remember from the equation. The dependent variable was recall accuracy of the equation elements (structural elements, e.g., parentheses; arithmetic operators; and the numerical values, i.e., the content of the equation).

Results

The results of Experiment 1 (see Figure 1) indicate that content and arithmetic operators were recalled better than structure at all levels of equation duration, Chi-square (2) = 13.4, *p*< 05. This suggests that content and operators may receive more initial processing time compared to structure. This finding does not support the hypothesis that people create a schema focused on structural elements.

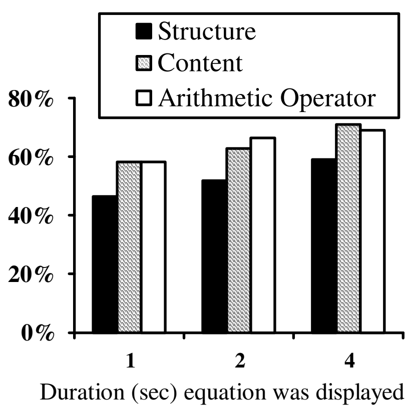

Figure 1. Percent of equation recalled correctly as a function of duration that the equation was displayed and the type of element in the equation.

A detailed analysis of the equations that participants produced during recall found that they tended to recall equation elements in a connected sequence or block, starting with the leftmost items of the equations (see Figure 2), Chi-square (7) = 142.9, *p*<.0001. This finding is consistent with the hypothesis that people read equations from left to right. Also, in all presentation conditions, participants tended to make more recall errors for the parentheses than for numbers or operators. This finding suggests that parentheses are not processed in the same way as numbers or operators, in that they serve a structural, not a semantic, role. That is to say, parenthesis help the reader navigate through the equation and identify how to organize it, but do not provide meaning. As a consequence, they appear to be easier to forget.

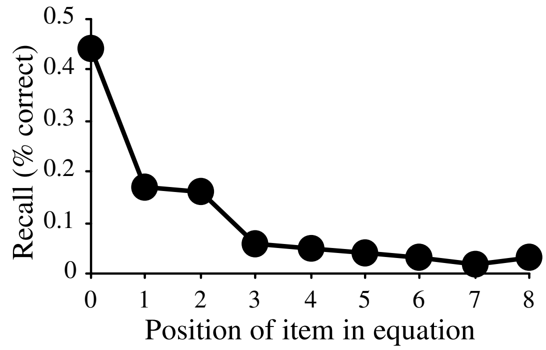

Figure 2. Mean percentage correct recall as a functio of the position of the item in an equation.

EXPERIMENT 2: EQUATION PREVIEWS

Experiment 2 further examined the hypothesis that an equation is read with an initial focus on the structure, after which, numerical content elements are filled in. We predicted that if people read the structural and arithmetic operators first in an equation, then a preview of structural and arithmetic operators before the whole equation should result in the shortest reaction time for solving an equation.

Method

Seventeen participants solved 36 computer-displayed equations -- nine each with one of four types of previews -- no preview, a preview of the entire equation, a preview of the parentheses and arithmetic operators, and a preview of the numbers in their position in the whole equation (see Figure 2). Each preview had a 2-sec duration, followed by the entire equation, which was displayed until the participant responded with the solution to the equation.

Structure preview	**Content preview**
(-) +	14 9 5 5

Figure 3. Examples of two of the preview types used in Experiment 2.

Results

The structural preview produced the longest overall solution times (main effect of Preview, $F(3,48) = 6.19$, $p = .0012$). Interestingly, with simple equations, all preview types resulted in faster equation solution times than the no preview condition, whereas, with complex equations, the structural preview slowed equation solution times compared to no preview and the other preview types (Preview x Complexity, $F(3,48) = 5.80$, $p = .0018$. The results for Experiment 2 indicate that previewing structural information slowed equation solution, contrary to the prediction from the hypothesis that when people read an equation, they perform an initial scan for structure.

EXPERIMENT 3: EQUATION CHUNKING

Experiment 3 investigated the hypothesis that, when they read an equation, people (1) decompose the equation into chunks defined by expressions contained in parentheses, (2) solve the expression, and (3) store the outcome in memory. We used a recognition task to study this hypothesis -- participants read a target equation, and then were given different recognition test probes. The probes were either identical to the target equation or differed from the target in terms of the expression within the parentheses or in terms of the equation outside of the expression. In some cases, the changes in elements in the probe resulted in the same solution as the target equation, whereas in other cases the target equation and the probe had different solutions.

Method

Each of 10 participants received a total of 72 trials. The procedure for each trial began with the presentation of a *target* equation and participants solved the equation mentally, at which point, they typed the answer into a response field on the computer screen. Then, a distractor task displayed a randomized starting number (e.g., 187) that was used as a starting point for the participant to mentally count backwards by 3 until the computer made a "beep" sound which signaled the participant to stop counting backwards and type the last number reached into a box on the computer screen. The duration of time from the presentation of the starting number until the computer beeped was a randomized time ranging from 3 to 15 seconds. Following the distractor task, the computer displayed a recognition test probe (an equation). The participant was asked whether the test equation was the same as the earlier target equation or not, and to rate the confidence of their response (on a 4-point scale from not sure to very sure). By combining the recognition response with the rating, each trial resulted in an eight-point response from "No, very sure" (0) to "Yes, very sure" (7).

An R-Index value (Brown, 1974), a bias-free measure of discrimination, was computed from the same/different response and the confidence rating by examining the overlap of the distributions of the 7-point ratings for the Identical condition with the different conditions. Perfect discrimination performance (i.e., no overlap between distributions) would result in an R-Index of 1.00, where as no discrimination (i.e., complete overlap between distributions) would result in an R-Index of .50.

On 36 of the trials, the test probe was identical to the target equation, and on 36 of the trials, the test probe differed from the target, with systematic differences among the probes. Table 1 provides examples of the seven versions of the test probe.

Results

Table 1 shows the results of Experiment 3. Participants responded that they recognized the test probes with proportions of "yes" responses among the six different probe conditions that differed significantly, Chi Square (5, N = 4500) = 269.5, $p < .0001$, with recognition scores that differed significantly as a function of the type of recognition probe, $F(6,54) = 20.41$, $p < .0001$, and with R-Indices that differed across the six "different" conditions.

Table 1. Versions of the test probe (with differences between the probe and Target equations in italics) and the proportion of "Yes" recognition responses, mean recognition score, and R-Index as a function of condition in Experiment 3.

Condition	Target Equation	Recognition (test) Probe	% Yes	Mean Score	R-Index
Identical	7(8/4)+1-12 = 3	7(8/4)+1-12 = 3	89	5.83	-----
Different numbers, same outcome in chunk	3(10/5)+5-8 = 3	3(*4*/*2*)+5-8 = 3	58	4.38	.58
Different numbers, different outcome in chunk	4(9/3)+6-10= 8	4(*8*/*2*)+6-10= *12*	33	2.45	.73
One different number, different outcome in chunk	6(12/6)+3-5= 10	6(12/*2*)+3-5= *34*	23	2.85	.79
Different numbers outside of chunk, same outcome	7(6/2)+4-15= 10	*5*(6/2)+*6*-*11*= 10	28	2.82	.79
Different numbers outside of chunk, different outcome	5(8/4)+6-12 = 3	*3*(8/4)+*4*-*6* = *4*	24	1.21	.80
One different number outside of chunk, different outcome	3(10/5)+1-2= 5	*6*(10/5)+1-2= *11*	34	2.58	.78

The identical recognition probe had significantly higher recognition scores than all other probe types. The probe that had two changes inside of the parentheses, but the same outcome within the parentheses produced a significantly higher recognition score than any of the other probes that differed from the target (Newman-Keuls, p's<.05). The lower R-Index for the probe with two changes, but the same outcome inside the parentheses, than for all of the other probes that differed from the target indicates that participants did good job discriminating the correct and incorrect probes, except on those trials with the difference in the parentheses.

Note that participants had higher recognition scores for that probe than for the probe that had only one different number within the parentheses (Newman-Keuls, p<.05). One possibility is that participants simply remembered the outcome of the whole equation and recognized when the outcome differed and when it was the same. However, the probe that had two changes outside of the parentheses with the same outcome also had the same outcome of the whole equation and it produced a proportion of "Yes" responses of only .28, a mean recognition score of -1.82 (significantly lower than both the identical probe and the probe with two differences in the parentheses, same outcome, Newman-Keuls, p<.05), and an R-index of .80. Thus, the results support the hypothesis that when people read equations, they substitute the outcome of parenthetical expressions (i.e., chunks of the equation) for the original numbers in the equation.

STUDY: EYE MOVEMENTS

Recording eye movements has been a valuable technique for understanding the perceptual and cognitive processes involved in reading text (e.g., McConkie, 1983). Accordingly, we have begun to study how people's eyes move as they read equations. This study describes our initial foray into recording eye movements during equation reading.

Method

In an initial study, 3 participants (all psychology graduate students) received 42 equations displayed on a 19-in ViewSonic G810 Monitor. Participants were asked to solve the equation and to speak the answer aloud. The participants' eye movements were recorded using an EyeLink®II eye tracking system and a Dell Dimension 4550 computer.

Results and Discussion

These eye movement traces showed that the initial fixation occured at the left side of the equation, as we predicted. In addition, participants tended to fixate on numbers for a longer perior of time and more frequently than on operators or parentheses, with fixations on parentheses occurring rarely. This is also consistent with the previous findings. Surpriwsingly, participants' eye movements did not tend to move in a left-to-right manner to the extent that we had expected. Rather, backscans followed by forward saccades were very common, especially with complex parenthethetical expressions. We are continuing this line of research and are expanding it to examine participants with both high and lowlevels of math expertise.

GENERAL DISCUSSION

Taken as a whole, the research described in this report suggests that people (1) read equations from left to right, one element at a time, (2) process operators and numbers more than parentheses, (3) store the outcome from a parenthetical expression, and (4) engage in frequent backward scans. The process tracing study also indicated that people might scan the entire equation to set up a structural schema before they solve the equation, but the experiments and eye tracking study failed to support that observation.

Application of the results to the design of the Math Genie

The present research provided the following guidance for developing an auditory browser for equation reading: (1) read one element at a time, distinctively, from left to right; (2) permit users to go back to the start of the equation from any point in the equation, (3) permit users to return to self-defined points in the equation from any point in the equation, (4) permit users to substitute (but not completely replace) the expression within parentheses with its outcome, and (5) permit users to scan the equation to provide an overview of the structure of the equation. We believe that, although the fifth guideline is not consistent with all of the data, providing users with this capability would not disrupt their equation reading performance, even if it would not necessarily be beneficial.

The above guidelines have been implemented in experimental software, the Math Genie (Karshmer, Gupta, and Gillan, 2002). The Math Genie (see Figure 3 for an example screen), uses a hierarchical structure to group subsections of an equation into meaningful units for browsing. The Math Genie also offers refreshable Braille output and specialized video output for low vision users.

Conclusions

The research that is described above has been useful in identifying some of the perceptual and cognitive processes that underlie equation reading by sighted people. We believe that persons with visual disabilities may gain some benefit from having the ability to access the information in equations in the same way that sighted people can, but by no means should visually disabled equation readers be forced to access that information in only one way. Universal access and universal usability as approaches to design should focus on providing users with as much flexibility as possible, while also providing guidance in the use of that flexibility.

The research in this paper, which grew out of the Mathematics Accessible to Visually Impaired Students (MAVIS), and its application in the design of the Math Genie can serve as a model for the ways in which different disciplines can interact in productive ways, while staying true to their disciplinary foci. The cognitive/perceptual psychologists on the project have been able to do research on how people read equations. The results of that research have then been taken up by the computer scientists who have been able to address issues of interest to them. The mathematicians involved in the project will ultimately have tools that will assist in math education. We also hope that, as the products of the research begin to be used, we will receive feedback from the users. That feedback will complete the first iteration in the design cycle and we will begin (1) to formulate new hypotheses about how people read equations, with a greater focus on the unique characteristics of how blind people read equations and (2) to improve on the usability and functionality of the Math Genie.

Acknowledgments

The work reported in this paper was funded under a grant from the U.S. National Science Foundation (HRD 9800209). The research described in Experiments 1 and 2 were part of the McNair Scholarship Program project of the first author.

References

Brown, J. (1974). Recognition memory assessed by rating and ranking. *British Journal of Psychology*, *65*, 13 - 22.

Karshmer, A. Gupta, G., and Gillan, D. J. (2002). Architecting an auditory browser for navigating mathematical expressions. In K. Miesenberger, J. Klaus, and W. L. Zagler (Eds.) *Proceedings of the International Conference on Computers Helping People with Special Needs: Lecture Notes in Computer Science 2398*. (pp. 477 – 485). Berlin: Springer-Verlag.

McConkie, G.W. (1983). Eye movements and perception during reading. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes*. (pp. 65 – 96). New York: Academic Press.

Nisbett, RE & Wilson, TD (1977). Telling more than we know: verbal reports on mental processes. *Psychological Review*,*84*, 231-229.

Smyth, M. M., Collins, A. F., Morris, P. E., and Levy, P. (1994). *Cognition in action.* Hove: Lawrence Erlbaum Associates Ltd.

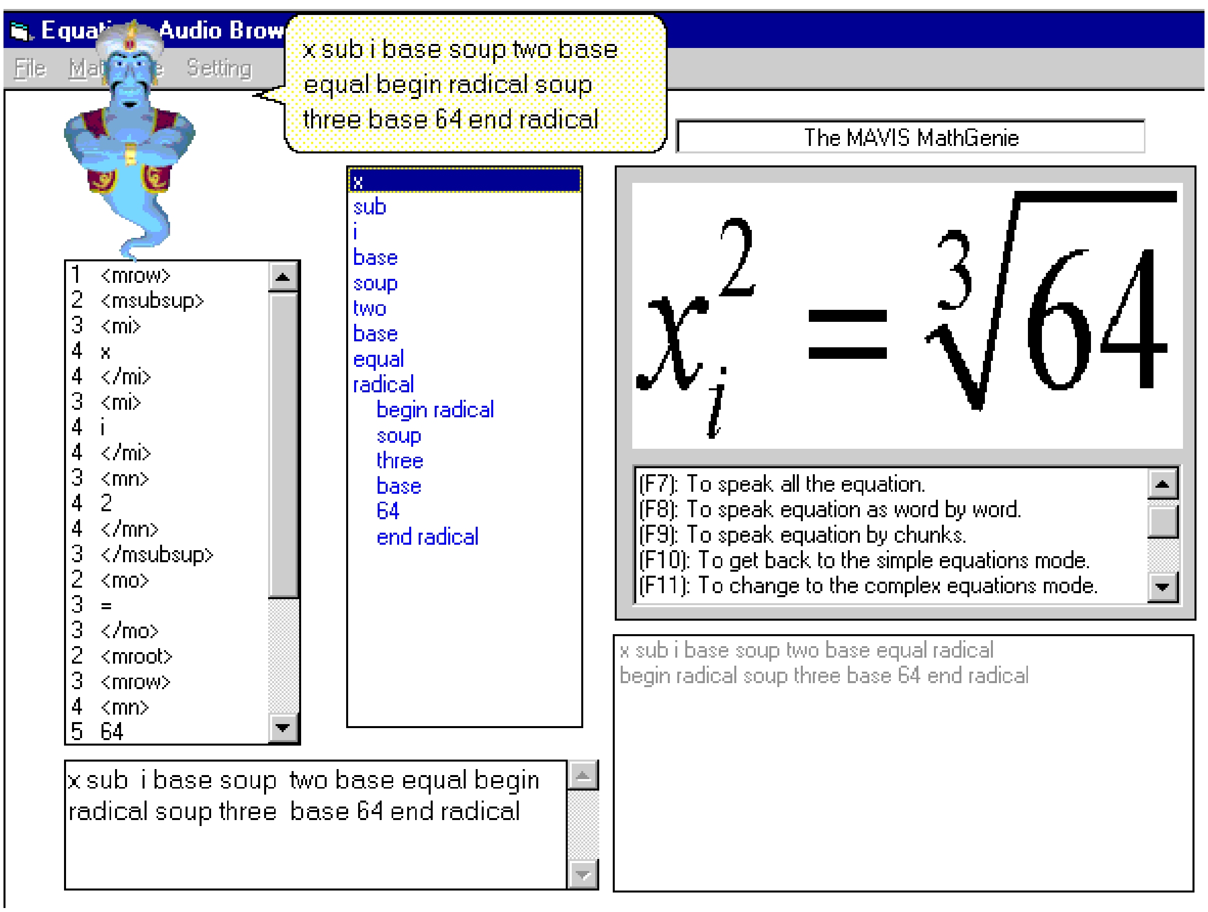

Figure 3. A sample screen from the MAVIS Math Genie.