

**UNDERSTANDING HOW VISUALLY IMPAIRED STUDENTS
DEMONSTRATE GRAPH LITERACY WITH ACCESSIBLE
AUDITORY GRAPHS**

A Thesis
Presented to
The Academic Faculty

by

Benjamin K. Davison

In Partial Fulfillment
of the Requirements for the Degree
Human-Centered Computing Doctor of Philosophy in the
School of Interactive Computing

Georgia Institute of Technology
October 2012

UNDERSTANDING HOW VISUALLY IMPAIRED STUDENTS
DEMONSTRATE GRAPH LITERACY WITH ACCESSIBLE
AUDITORY GRAPHS

Approved by:

Dr. Bruce N. Walker, Advisor
School of Psychology
Georgia Institute of Technology

Dr. W. Keith Edwards
School of Interactive Computing
Georgia Institute of Technology

Dr. John Stasko
School of Interactive Computing
Georgia Institute of Technology

Dr. Richard Catrambone
School of Psychology
Georgia Institute of Technology

Dr. Tony Stockman
School of Computer Science
Queen Mary University London

Date Approved:

TABLE OF CONTENTS

| | |
|--|-----|
| LIST OF TABLES | vii |
| LIST OF FIGURES | ix |
| GLOSSARY | 1 |
| I INTRODUCTION | 1 |
| 1.1 Overview of Graphs | 4 |
| 1.2 Thesis | 9 |
| 1.3 Document Overview | 13 |
| II BACKGROUND | 15 |
| 2.1 Graph Literacy | 16 |
| 2.1.1 Anscombe’s Quartet | 16 |
| 2.1.2 Graphs and Efficiency | 18 |
| 2.1.3 Discussion | 20 |
| 2.2 Auditory Graph Fundamentals | 20 |
| 2.2.1 Acoustics and Psychoacoustics | 21 |
| 2.2.2 Mapping | 24 |
| 2.2.3 Context | 25 |
| 2.2.4 Trend Analysis | 29 |
| 2.2.5 Point Estimation | 29 |
| 2.2.6 Conclusion | 30 |
| 2.3 Auditory Graphs as Educational Technology | 30 |
| 2.3.1 Sound Graphs: Successful Trend Analysis | 31 |
| 2.3.2 Plotting points with the Integrated Communication to Draw . . . | 32 |
| 2.3.3 Upson’s Education Evaluation | 35 |
| 2.3.4 Audio-Haptic Graphs in MULTIVIS and related projects | 39 |
| 2.3.5 General Discussion | 46 |
| 2.4 Current Practice: Curriculum, Graph Literacy, and Visually Impaired Stu- | |
| dents | 47 |
| 2.4.1 Understanding Curriculum | 47 |

| | | |
|-------|---|-----|
| 2.4.2 | Common Core Standards | 54 |
| 2.4.3 | Testing Accommodations | 56 |
| 2.4.4 | Graphs and Number Lines for Visually Impaired Students | 57 |
| 2.4.5 | Discussion | 62 |
| 2.5 | Conclusion | 64 |
| III | TECHNOLOGY | 65 |
| 3.1 | AGM: Representing Graph Data, Context, and Presentation | 65 |
| 3.1.1 | Storing Graph Data | 66 |
| 3.1.2 | A Window of Opportunity | 68 |
| 3.1.3 | Generic Settings for Displays | 71 |
| 3.1.4 | Limitations | 72 |
| 3.2 | AGFC!: Converting between the Model and Other Formats | 72 |
| 3.3 | AGE!: Combining Model and Conversion | 78 |
| 3.3.1 | Discussion | 79 |
| 3.3.2 | Future Work | 81 |
| 3.4 | AudioFitts: Discovering Active Point Estimation | 83 |
| 3.4.1 | A Blind Student Learning to Graph | 83 |
| 3.4.2 | Design | 86 |
| 3.5 | Navy: Training in Active Point Estimation | 90 |
| 3.6 | GNIE: Active Point Estimation In Classrooms | 92 |
| 3.6.1 | Future Work | 93 |
| 3.7 | General Discussion | 93 |
| 3.8 | Conclusion | 94 |
| IV | PHASE 1: POINT ESTIMATION WITH AN AUDITORY DISPLAY | 95 |
| 4.1 | Study 1: The effect of input device on target speed for visually impaired participants. | 95 |
| 4.1.1 | Study Design | 99 |
| 4.1.2 | Results | 107 |
| 4.1.3 | Discussion | 111 |
| 4.2 | Study 2: Determine the effect of input device and presentation format on target speed for sighted participants. | 115 |

| | | |
|-------|--|-----|
| 4.2.1 | Study Design | 116 |
| 4.2.2 | Results | 118 |
| 4.2.3 | Discussion | 121 |
| 4.3 | Study 3: Determine the effect of auditory scaling and peaking on target speed and accuracy. | 122 |
| 4.3.1 | Study Design | 124 |
| 4.3.2 | Results | 126 |
| 4.3.3 | Discussion | 127 |
| 4.4 | Study 4: Determine the effect of mapping type and auditory scaling on target speed and accuracy. | 129 |
| 4.4.1 | Study Design | 129 |
| 4.4.2 | Results | 130 |
| 4.4.3 | Discussion | 131 |
| 4.5 | General Discussion | 131 |
| V | PHASE 2: FROM GRAPHING STANDARDS TO A GRAPHING SYSTEM . | 133 |
| 5.1 | Standards: Identify education standards used in graphing. | 133 |
| 5.1.1 | Study Design | 134 |
| 5.1.2 | Results | 135 |
| 5.1.3 | Discussion | 135 |
| 5.2 | Finding Relevant Questions | 138 |
| 5.2.1 | Study Design | 140 |
| 5.2.2 | Results | 142 |
| 5.2.3 | Discussion | 142 |
| 5.3 | Identifying the Steps Needed to Answer the Graphing Questions | 143 |
| 5.3.1 | Study Design | 143 |
| 5.3.2 | Results | 144 |
| 5.3.3 | Discussion | 145 |
| 5.4 | Reconstruct the Steps in an Auditory Graphing Technology | 146 |
| 5.5 | General Discussion | 147 |
| 5.6 | Conclusion | 148 |

| | | |
|------------|--|-----|
| VI | PHASE 3: EVALUATION IN CLASSROOM AND TESTING SCENARIOS . | 149 |
| 6.1 | Classroom Evaluation | 149 |
| 6.1.1 | Study Design | 149 |
| 6.1.2 | Pilot Study | 152 |
| 6.1.3 | Results | 158 |
| 6.1.4 | Discussion | 163 |
| 6.2 | Examination Performance | 163 |
| 6.2.1 | Study Design | 164 |
| 6.2.2 | Pilot Study | 171 |
| 6.2.3 | Results | 176 |
| 6.2.4 | Discussion | 179 |
| 6.3 | General Discussion | 183 |
| APPENDIX A | COMMON CORE STANDARDS | 184 |
| APPENDIX B | GRAPHING QUESTIONS | 192 |
| APPENDIX C | TASK ANALYSIS | 202 |
| APPENDIX D | SURVEY INSTRUMENTS | 203 |

LIST OF TABLES

| | | |
|----|---|-----|
| 1 | Research Questions | 12 |
| 2 | Contributions | 13 |
| 3 | Anscombe’s quartet, data table. See figure 5 for the graph of this data. . . | 17 |
| 4 | Key features of Figure 12, as defined by [?]. Note the emphasis on trends. | 41 |
| 5 | Graph component aspects of evaluation, throughout the 2000-2010 research in multimodal graphs in the Brewster lab. Table is broken by line, point, and hybrid evaluations. | 43 |
| 6 | Other aspects of evaluation, throughout the 2000-2010 research in multimodal graphs in the Brewster lab. | 44 |
| 7 | A comparison of requirements from Form 1 (9th grade) mathematics, section 19.0.0, COORDINATES AND GRAPHS. See [? ?]. Notice how the blind syllabus emphasizes <i>exploring</i> graphs, while the standard syllabus emphasizes <i>exploring and authoring</i> graphs. | 62 |
| 8 | An overview of the AGFC! (AGFC!) abstractors and reifiers. There are five abstractors and 14 reifiers, four of which can do both. | 74 |
| 9 | Graphing behaviors for Vivian, a sighted student using paper, Cindy, a blind student using the Sonification Sandbox, and proposed behavior for a blind student. | 85 |
| 10 | Variables available for manipulation in AudioFitts. Most of these variables are described further in the specific experiments in Phase 1 (Chapter 4). . . | 89 |
| 11 | Scores for the Navy game. | 92 |
| 12 | The four studies in phase 1. These establish how to use sonification for point estimation. | 95 |
| 13 | The effect on the cursor position from pressing the 10-key number pad. The ‘none’ column of numbers have no effect on movement. | 96 |
| 14 | Study 1 statistical tests conducted. Effect size is for μ_p^2 (partial eta squared) for the Repeated Measures ANOVA, and Cohen’s d for Student’s t-test. Contrasts are linear. Horizontal lines indicate families of tests. | 107 |
| 15 | Study 2 significance tests conducted. All results are significant, except for visuals-combined. | 118 |
| 16 | Descriptive statistics for speed for the interaction device and sensory condition. The mean values are first, followed by the standard deviation in parentheses. | 120 |

| | | |
|----|--|-----|
| 17 | The different scalings evaluated. The top row is the distance from the target, and the following rows show the MIDI note for each scaling. The actual MIDI range will be about 40-80. For the formulas, the MIDI note is bounded within the range. | 123 |
| 18 | Accuracy of peak and auditory scaling conditions for Study 3. | 126 |
| 19 | Speed of peak and auditory scaling conditions for Study 3. In seconds. . . . | 127 |
| 20 | Study 3 results. | 130 |
| 21 | Standards results, parts 1 and 2: The graphing standards that include the text “coordinate”, “graph”, and “number line”. Standards excluded in part 2 have a strikethrough. | 136 |
| 22 | Standards results, part 3: The final graphing standards, based on CCS! (CCS!) Mathematics Grade 6. | 137 |
| 23 | The component behaviors that need to be supported by the system, in order to allow students to answer all of the graphing questions. “Estimate Values” means estimating the current position, between two tick marks with known values. “Filled Region” means setting a region larger than a point that represents the values; this is for number lines only, such as a number line of $x > 3$ | 144 |
| 24 | The behavioral components’ relationship to the graphing standards. | 145 |
| 25 | Results from the Navy Game experiment. ”D” stands for ”day”, ”R” stands for ”round”, and ”Cum” stands for ”Cumulative” (the sum of all rounds for the day). There were significant improvements in 5 out of six of the tests, showing clear signs of students improving on navigating toward targets. . . | 159 |
| 26 | Scores from the exit survey during the summer camp (n=19). The survey asked participants to ”answer with a scale of 1 to 5, where 1 is strongly disagree, 2 is somewhat disagree, 3 is neutral, 4 is agree, and 5 is strongly agree.” | 162 |

LIST OF FIGURES

| | | |
|----|---|----|
| 1 | A mainstream high school classroom in Georgia. Computing is everywhere: in the desks, on the smartboard, and shown on the projector. The only blind student sits at the back of the room with her vision teacher and tactile graphics. | 3 |
| 2 | Two sample graphs. From [? , page 27]. | 4 |
| 3 | Percent of second-level curriculum requirements in the GPS! (GPS!) requiring graphing. Based on the presence of words “number line”, “coordinate”, and “graph” in the deepest descriptive level of [? ? ? ? ? ? ?]. Results are for each grade; high school mathematics is by course name. Every grade has graphing requirements in the standards; 9 of 13 classes have graphs and number lines as over 15% of the standards. | 7 |
| 4 | An example of tactile graphics. | 7 |
| 5 | Anscombe’s quartet of visual graphs [?]. The original data is in table 3. All of these graphs have the same x and y mean, variance, correlation, and line of fit. However, the graphs are visually distinct, with different types of trends. The top left shows a moderate linear correlation between x and y, slowly increasing. The top right shows a parabola-like curve facing downward. The bottom left is a strong linear correlation between x and y, with a single outlier. The bottom right is a vertical line at x=7, with an outlier. | 18 |
| 6 | The relationship between frequency and pitch, as measured by Stephens et al. [?]. As noted in [?], the pitch changes quickly for low frequencies and slower for high frequencies. Note the frequency scale is logarithmic. | 23 |
| 7 | The graph of $f(x) = 0.25x^2$, on the domain of $-0.5 \leq x \leq 4.1$ with and without context. In both cases, it is possible to see where the graph is increasing, decreasing, and flat. The function also looks like a parabola. However, the first image cannot be used to determine intercept or any of the points. | 25 |
| 8 | A portion of figure 4 from Smith and Walker [?]. Note that the point estimation of each participant’s recreation of the black line (in gray) is close, but probably varies from the correct answer more than what a mathematics teacher would judge to be correct. | 27 |
| 9 | Grids in the Integrated Communication 2 Draw. From [?] | 33 |
| 10 | Upson’s Sound Grid, featuring an interactive, two-dimensional canvas, and settings for auditory graph playback. From [?]. | 37 |
| 11 | Auditory display used in Ramlooll et al. [?]. Pitch is mapped to y values. Volume and pan are mapped in a way to simulate what would be heard from a sound source if the listener were standing at the origin, facing the positive direction of the x axis. | 41 |

| | | |
|----|--|----|
| 12 | An example of graphs from a 2003 evaluation, “Drawing by Ear” [?]. The leftmost graph is the original data. The middle and rightmost graphs are participant drawings of the graph, after listening to an auditory graph of the data. Note the differences in the trends and lack of verbal context. In one aspect, trend analysis, the drawings look similar to the original. However, without points, it is impossible to know the scale of the resulting graph. Participants were not given verbal cues for the task, which probably hindered their construction of the graphs in the first place, and led to minor errors in the trends. | 42 |
| 13 | The Graphic Aid for Mathematics. Push pins represent points, including the origin. Rubber bands represent lines, including axes. Note that this tool allows easy graph plotting and exploration, visually and with touch. However, it lacks an easy way to label. It is also not an electronic format, nor can it easily be scanned or converted into an electronic format. | 60 |
| 14 | Two graphics from Kenya. Note how the figures can be explored but may be difficult to create by a blind student. | 63 |
| 15 | Code for a graph with two lines. The first line has 4 points, along the path of $y = 2 * x - 3$. The second line is simply represented by the formula $y = -x + 1$ | 67 |
| 16 | A Yahoo Finance graph of the S&P 500 value in 1991 and 2012. The trend of the drop and recovery is similar, but the magnitude is different (60 compared to 150). | 69 |
| 17 | The S&P 500 values, 1955-2012. The overall pattern is visible, but smaller patterns found in Figure 16 can no longer be perceived. Figures from Yahoo Finance. | 70 |
| 18 | Setting the Window properties on a graph. | 71 |
| 19 | Setting MIDI Rendering settings for a graph. When renderings are incomplete or completely missing, the display reverts to default settings. | 72 |
| 20 | An overview of the AGE! (AGE!). Artifacts can be converted to an AGM! (AGM!) model by the AGFC! (blue lines), a process called abstraction. AGM! models can be converted by the AGFC! into artifacts (red lines), a process called reification. Some artifacts can be converted both ways such as the Saved Model. | 73 |
| 21 | The Sonification Sandbox software features opportunities to explore data in multiple formats simultaneously. This screenshot shows the user interface that features a visual graph, an auditory graph, and text description of the graph | 75 |
| 22 | Code for importing a saved graph file (abstraction), then exporting an image (reification). | 76 |
| 23 | Multiple reifiers can convert various AGM! graphs in a manner of milliseconds. The “Mass” demonstration program is shown above; it shows threaded conversions of a model into various formats. | 77 |

| | | |
|----|--|-----|
| 24 | A simplified, web-based version of the Sonification Sandbox. Enter in a formula to see a visual graph, and auditory graph, and a text description. As of 2012, this is a living website available at http://sonify.psych.gatech.edu/graph . | 80 |
| 25 | The leftmost image and the center image are two separate visual graphs. In a software program, these could represent graph models that can be manipulated. Dragging them together could combine the data, making a graph like that in the rightmost image. It may also be possible to separate the data apart again. | 82 |
| 26 | Visuals from an AudioFitts Experiment. The number line has tick marks, which can be accessed by moving the mouse left or right. | 87 |
| 27 | The Navy game used to encourage point estimation practice. The game is full screen, with a number range of 0 to 9. Players move the mouse to listen for ships and numbers, and press the number that is near the ships. In this example, pressing “7” would destroy 3 ships and give six points, while pressing “5” would not hit any ships and cost 1 point. | 91 |
| 28 | The GNIE program. | 93 |
| 29 | The mouse and the right side of the keyboard used in the experiment. Note the additional keys around the number keys, and the horizontal tactile mark on the 5 key. | 100 |
| 30 | The relationship between the distance to the target edge and the MIDI pitch, using the logarithmic scale. When inside the target, the user hears an alert. See Equation 5 and Table 17 for examples on the effect of the logarithmic mapping. | 103 |
| 31 | Target width, target distance range, and keyboard mean movement time. | 108 |
| 32 | Target width, target distance range, and mouse mean movement time. | 109 |
| 33 | Index of difficulty and movement time for the mouse and the keyboard. | 110 |
| 34 | The visuals for Study 2. The vertical gray line is the target. The vertical red line is the cursor. The black edges are the sides of the screen. | 116 |
| 35 | Box plots of the movement time, based on the sense and interaction device. | 119 |
| 36 | A diagram showing different scalings and peaks. Boolean scaling is vertical, following the black box down. LinearShort scaling is red, dropping quickly but a change for each pixel distance. LinearLong is blue, with a MIDI note drop every 3 pixels and with changes 3 times further than LinearShort. Logarithmic, green, drops the most at first, but has small changes hundreds of pixels away. Peaking is shown in orange. If peaking is on, MIDI notes change when the user’s selection pixel is within the target; if off, the MIDI note does not change. | 122 |
| 37 | The number line used in study 3. | 125 |

| | | |
|----|---|-----|
| 38 | A diagram of how the standards are pushed into the examination and the textbook. Both the tests and the textbook are inspired by the standards. However, the textbook has a different creator (publishers), and a particular graphing question may be less aligned with the standards than most graphing questions on a test. | 139 |
| 39 | Page 22 from the Mathematics 1 textbook [?]. All six problems on this page are graphing problems, and a red box is superimposed around each one. Based on the criteria for graphing question creation, problem 28 is the only one that could be used immediately; all of the other ones would have to be simplified. | 141 |
| 40 | The Max/MSP receiver. Each group is a single channel. Two channels are shown, 16 are available (one for each MIDI track), but only a maximum of four were used. | 146 |
| 41 | Students training on point estimation on a number line. | 154 |
| 42 | A student using GNIE to solve a graphing problem. | 155 |
| 43 | A teacher working with a student on paper graphs. | 156 |
| 44 | Mean scores of the Navy Game as a percent of the total possible score for each round. Note that the scores can be below zero. | 159 |
| 45 | Students worked in pairs of two, one of which used auditory graphs on the computer and one who used visual or tactile graphics. The teacher talked with the students, too. | 160 |
| 46 | Low vision students could magnify GNIE! (GNIE!) | 161 |
| 47 | This student is using a combination of visual graphs and tactile graphics to answer the questions. | 161 |
| 48 | The instructions for the test were in three formats. From left, the images show visual graphs, tactile graphics, and auditory graphs. | 168 |
| 49 | the graphs for the tests were in three formats. From left, the images show visual graphs, tactile graphics, and auditory graphs. | 168 |
| 50 | The graphs for the tests were in three formats. From left, the images show visual graphs, tactile graphics, and auditory graphs. | 169 |
| 51 | Identifying the point that is a reflection of (2,3) across the X and Y axes. Graphing Standard 7. | 173 |
| 52 | Finding the coordinates of point A. Graphing Standard 9. | 173 |
| 53 | Plotting the point (4,1) on the coordinate plane and identify the quadrant. Graphing Standard 6. | 174 |
| 54 | Plotting the reflection of (5, -4) over the x axis. Graphing Standard 7. . . . | 174 |
| 55 | Test scores by display condition. | 176 |

| | | |
|----|---|-----|
| 56 | Test score averages for each standard, comparing large print ("paper") to the GNIE ("audio") scores. The center green line is $y=1x + 0$, the theoretical fit line for the scores. | 178 |
| 57 | Test score averages for each standard, comparing braille ("paper") to the GNIE ("audio") scores. The center green line is $y=1x + 0$, the theoretical fit line for the scores. | 179 |
| 58 | Test score averages for each standard, comparing braille ("paper") to the GNIE ("audio") scores, and omitting questions for standards 5 and 7. The center green line is $y=1x + 0$, the theoretical fit line for the scores. | 180 |
| 59 | Histograms for the scores from low vision users, showing large print (top) and GNIE! (bottom) conditions. | 181 |
| 60 | Histograms for the scores from tactile graphics users, showing tactile (top) and GNIE! (bottom) conditions. | 182 |

CHAPTER I

INTRODUCTION

In their youngest years, children are encouraged to explore the diverse world. They use their eyes, ears, hands and feet to learn how to hold, how to walk, and how to build. The child's activities add a level of understanding how to do things that goes beyond any instructions. After awhile, children can ride a bicycle or read with minimal thinking. In later years, science provides lab and nature activities to deeply understand the world around us. Learning is an active process.

Mathematics is also active. Through a set of formal rules, the mind and body are asked to play with information so that insight can be found. Adding numbers, reducing fractions, and finding the area under the curve are tools learned over many years. Perceptual processes are a key component as well. Many students can use spatial layout to solve problems in their math activities. Fractions, geometry, and coordinate algebra all leverage the human visual system to quickly understand numbers. Coordinate algebra is particularly important in K-12 education, accounting for about 20% of the learning standards for mathematics.

Coordinate graphs and number lines present data in a form that can be simpler to interpret than tables, summary statistics, or text descriptions. A small visualization makes it possible to interpret the massive data of the stock market, such as overall change, fluctuations throughout the day, trading volume, and comparisons to other markets. Alternative presentation formats are possible, but have their drawbacks. Statistics such as mean, standard deviation, and fit lines can be the same for data that are clearly different when seen [?].

Graphs appear when the communicative intent is to show data relationships in a compact and sophisticated form. Jobs that require data analysis, including careers in science, technology, engineering, and mathematics (STEM), often use graphs. In other words, a job requirement of many STEM careers is to interpret graphs. Anyone who cannot interpret

graphs may be at a disadvantage in these fields. As a result, learning to use graphs should, and generally does, begin during a student's basic math education.

Educators have made coordinate graphs a core component of primary and secondary mathematics education. The emphasis of graphic literacy, or graphicacy [?], can be seen in many education standards. In Georgia, number lines and graphs are a part of the mathematics education standards in every course from Kindergarten through grade 12[? ? ? ? ?]. Even as young as five years old, students are learning the basics of graph literacy.

Unfortunately, the visual nature of a typical number line or coordinate graph presents a major hurdle for those with visual impairment. Due to limited options for accessible graphs in education and employment, people with vision impairment are at a disadvantage when trying to complete mathematics homework or work in STEM careers. And this handicap appears to be growing: the current practice of tactile graphics is increasingly incompatible with a blind student's education for two reasons. First, visually impaired students are increasingly placed in mainstream schools, in classes with their sighted peers. According to the Annual Report from the American Printing House for the Blind, an organization with an annual goal of knowing how many visually impaired students there are for each state, 83% of legally blind students attend mainstream schools, while 9% attend residential schools for the blind [?]. While mainstream schools may provide talented teachers and positive social opportunities for the student, it is more difficult to find sufficient human resources and equipment for complicated materials like graphs [? ? ? ?]. Second, computers are becoming more common in all classrooms. Some education activities, such as graphing, are difficult or impossible to complete using currently available computer software. As classrooms become more computerized, blind students are becoming *less* equipped to interpret and create and interpret graphs along side their peers.

Currently, while sighted students are moving to computers in STEM classes, the blind student is left with tactile graphics. Figure 1 is a photograph taken in a mainstream Georgia ninth grade mathematics classroom in October 2011. The front of the class has two computers, one connected to a smartboard and one to a projector. Each desk is embedded



Figure 1: A mainstream high school classroom in Georgia. Computing is everywhere: in the desks, on the smartboard, and shown on the projector. The only blind student sits at the back of the room with her vision teacher and tactile graphics.

with a desktop computer designed to work with a software learning system. The blind student, however, cannot gain access to the large portion of graphs and figures described visually. Instead, she works with tactile graphics in the back of the room with her vision teacher. In addition to class time, the student spends 90 minutes per school day with the vision teacher, often working on mathematics content. In terms of human resources, this student is well-served; most visually impaired students see their vision teacher once per week, and mostly cover non-content material. However, even with this support, the classroom teachers are still straining to get the blind student through the graphing parts of the course.

One solution to this digital divide in graphing accessibility in computers is to look for computerized forms of accessible graphs. On a standard desktop computer the natural alternative to visuals is audio. Along these lines, accessible auditory graphs research began 26 years ago with “Sound Graphs” [?]. Auditory graphs have promised to provide access to data, often without the use of any additional equipment. Auditory graphs can be used to detect trends, find patterns, and use context (e.g. [? ?]). And yet, with limited exceptions [?], auditory graphs have not been evaluated with visually impaired students in classroom

Find the slope of the line that passes through the points.

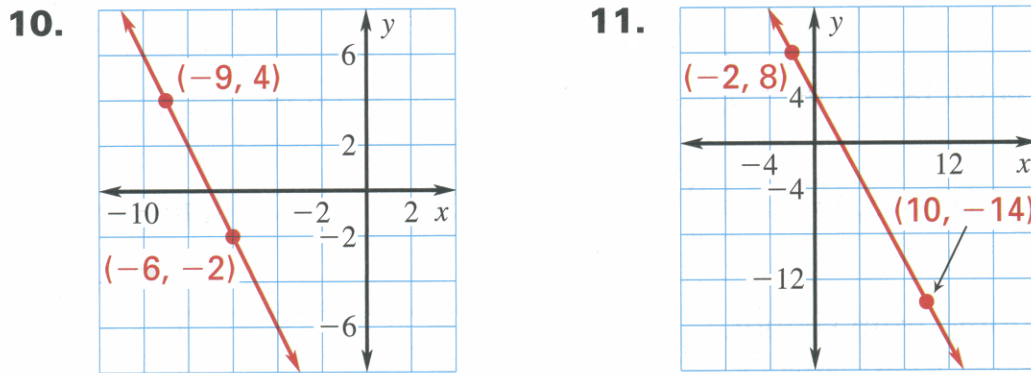


Figure 2: Two sample graphs. From [? , page 27].

or testing environments.

Auditory graphs are a potential solution to the graphs disability challenge. Blind students and their teachers may find auditory graphs as an acceptable alternative to visual and tactile graphs. Auditory graphs also can often be generated with standard desktop equipment. This dissertation proposes to demonstrate effective auditory graphs through a process of understanding Georgia K-12 graphing problems and evaluating alternatives to visual and tactile formats. It presents an auditory graph tool, the Graph and Number line Interaction and Exploration system (GNIE), as software to be used in a realistic learning environment for blind students. This chapter begins with a brief introduction to graphs, graphicacy, math education, graphs for the blind, and auditory graphs. It will then present the thesis, research questions, and contributions.

1.1 Overview of Graphs

What is a “graph”? People have used graphs for hundreds of years, and the meaning of “graph”, “map”, “chart”, “figure”, “information graphic”, and other terms can refer to several categories of artifacts; one dictionary’s definition of “graph” is as broad as “a written symbol for an idea, a sound, or a linguistic expression,” [?]. Before presenting a more suitable definition for this dissertation, this section describes how others have approached graphing, in terms of use, cognition, and alternatives.

In Georgia high school mathematics textbooks, a graph is a plot of two dimensions of data. Figure 2 has two examples. A coordinate graph displays a relationship between two variables, typically x and y . Books display a visual picture where pairs of number values, such as 8.2 and -7, are plotted along axes in a way that allows spatial position and distance to compare any values, or to get an overview of all of the values. The data values can be extracted from the graph¹, and the trend of the data is revealed, often much more easily than through data tables or statistics [? ?].

The graphs in Figure 2 contain many design elements that help a viewer interpret the information. One important observation of the graphs is that most of the ink used to print the graph represents context information such as tick marks, axes, and labels [?]. A repeating pattern of vertical and horizontal lines (blue in this example) creates a background grid, with each line representing a successive step in value. For the x and y axes, the graph shows a thick black line with an arrowhead on each side. Each axis is labeled, along with certain grid lines, next to where they cross the axis line. Along with this context, there are representations of the data, all printed in red: two points; ordered pair labels near the points; and a line intersecting the points. The data can be easily understood because of the medium the context provides.

This graph provides an efficient medium for interpreting data trends and values. Axis labels and grid lines provide a way to find the value of a line at a particular place. The “zero-line” axes in the grid have shifted left and up between problems 10 and 11, to best suit the graph being displayed. For example, problems 10 and 11 each have a point which cannot be displayed within the context provided in the other problem. Colors indicate functional differences, such as context and data. Sighted students can also see trends, and visually deduce certain qualities of the function, such as “the graph in problem 10 does not pass through the first quadrant, so there are no x/y pairs in the function that are both positive.” Both slopes clearly go downward from left to right, indicating a negative slope. Thus, an algebraic calculation of the slope must produce a negative number.

Yet, while useful, the graph hides certain information that could help interpretation.

¹The actual values are retained within a small error.

Only certain grid line labels are given along the axis. The grid lines have a step size greater than one. Three points are not placed at grid line intersections, making it more difficult to deduce the point values (without the given labels). It is not shown where the graph in problem 10 crosses the y axis. The problem of solving the slope is algebraic and really only needs the ordered pairs, not the graphs. But the graphs appear to give some insight about, for example, the expected slope. Graphs, then, appear to give some intuition about relationships, perhaps more easily than solely through data tables and formulas.

State and national education organizations list graphs as an important, continuing component in their mathematics learning requirements. For example the Georgia Performance Standards (GPS), the curriculum for K-12 students in the state of Georgia, has a set of requirements for each grade². Starting in first grade, the GPS makes explicit reference to graphs. By third grade, the curriculum states as a requirement: “construct and interpret line plot graphs” [?]. Figure 3 shows the percentage of a grade’s curriculum describe requirements for graphing. Note that the percentage of graphing topics in the curriculum increases in later grades, to about 20%. A similar emphasis on graphs in the mathematics standards can be seen in other curricula, including the interstate Common Core Standards (CCS), which will be used in almost all of the states in the United States by 2014 (including Georgia).

The most common graph alternative for blind students is tactile graphics. A tactile graphic is a surface with raised dots, lines, and regions. Figure 4 provides a few examples. Tactile graphics have been developed for over 200 years[?]. Tactile graphics are produced in a variety of ways, many of which require special skills and vision.

Many tactile graphics cannot be changed after production: they are indelible [?]. Although modern embossers can create tactile graphics from a computer, typical computer practice requires a sighted person to design a tactile graphic; these tools are not accessible for the blind. A dynamic tactile display (DTD) such as the Optacon[?] may overcome the indelibility problem. However, DTDs often require expensive equipment for a computer, up to \$100 for a single braille cell. The Optacon also had drawbacks[?], many of which affect

²In high school, the grade distinctions are replaced by class distinctions, such as Mathematics 1.

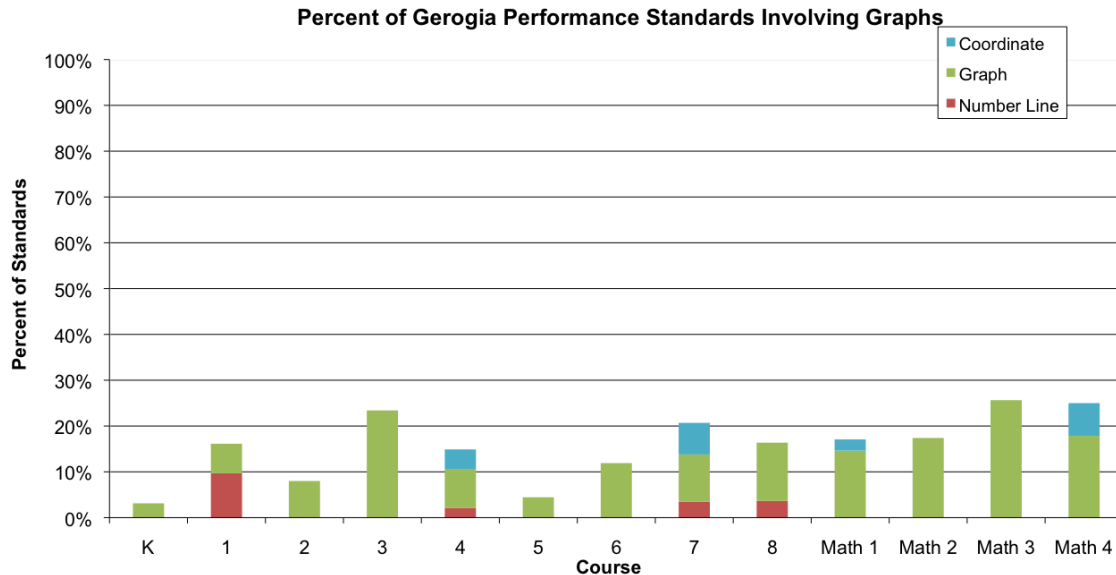


Figure 3: Percent of second-level curriculum requirements in the GPS requiring graphing. Based on the presence of words “number line”, “coordinate”, and “graph” in the deepest descriptive level of [? ? ? ? ? ?]. Results are for each grade; high school mathematics is by course name. Every grade has graphing requirements in the standards; 9 of 13 classes have graphs and number lines as over 15% of the standards.



Figure 4: An example of tactile graphics.

all modern DTDs. An alternative to DTD that may support a computerized infrastructure and blind student authorship is auditory graphs.

Bly[?] and Mansur and Blattner[?] introduced ways to map coordinate graph data into sound. Mansur and Blattner were concerned about blind people’s access to graphs. Their studies created sonifications, or mappings of data to non-speech audio, by mapping mapping of x-values to time, and y-values to pitch. This practice has been so effective it has been the standard sonification mapping for the past 25 years. Mansur and Blattner also framed the challenges of tactile graphics, including portability and requiring help from a sighted peer[?]. In my experience, many visually impaired students are introduced to tools that contain auditory graphs, such as the Audio Graphing Calculator³ or MathTrax⁴. These tools, however, are shown briefly, *after graphs have already been learned*. It appears that they are not used as a primary learning tool.

Robert Upson explored how auditory graphs could be used to learn graphs in this way. Upson taught sighted students how to use auditory graphs, and collected their test performance and opinions of auditory graphs [? ?]. Unfortunately, Upson did not find major improvements, and he didn’t work with blind students. Sighted students may not be motivated to use nontraditional forms of graphs. In addition, Upson was concerned about the general utility of auditory graphs. A more specific set of mathematics problems may highlight more advantages.

My work will extend Upson and related research in accessible math education, graphi-cacy, and auditory graphs in a number of ways. This work program of research begins with psychoacoustic studies of sonifications for point estimation. Then, a novel method called Standards, Questions, Answers, Reconstruct, and Evaluate (SQUARE) was used to ground the graphing problems on a particular set of requirements, the CCS for Mathematics, grade 6. The final phase used the resulting system in evaluations of testing performance and impact of auditory graphs in a classroom.

So what is a “graph”? Based on the work of others and the focus of this dissertation,

³The Audio Graphing Calculator is available through ViewPlus, <http://downloads.viewplus.com/software/AGC/>.

⁴MathTrax is available through NASA, <http://prime.jsc.nasa.gov/mathtrax/>.

Definition A *graph* is an interactive display that gives a person access to non-verbal relationships between parts of the data within and between dimensions. These data are mapped onto a context which uses a scale and a mapping for within each dimension. The context also includes verbal numerical data as a guide for point estimation.

1.2 Thesis

This dissertation is concerned with practical issues of accessible auditory graphs in a classroom environment. The definition of a graph informs the thesis in a number of ways. First and foremost, a graph is an interactive display. The interactivity of a graph comes from the availability of information. In many visual graphs, a mouse or finger is not needed for the interaction; the eyes can explore the data. For alternative formats, a tool may be needed to assist the user in exploring the graph, but the concept of user-driven exploration is critical. For, just as a text description could give a serial, already-interpreted version of the graph, an uncontrolled version of the graph would not lend itself for a user to mine the graph for particular meaning. The best graphs beg for an active participant.

Second, a graph is primarily non-verbal. The relationship between points on a graph is facilitated by perceptual processes such as closeness in space, color, or tone. This understanding of data is different than comparing numbers, and for large data sets, a graph can lend itself to finding patterns. While verbal components (labels) are a critical part of a graph, they are not the center of the perceptual process.

Third, a majority of information in a graph is its context. Facilitating point estimation, or the process of guessing a value, is as a fundamental component to a graph as enabling trend analysis, or discovering patterns in data.

Parts of the definition have been explored by others. Auditory graphs, or sonifications designed as alternatives to simple visual line graphs and scatterplots, have been known since the 1980s [? ?], and continue to drive active research. Walker, Smith, and Nees [? ?] (among others) have presented the benefits of context. However, this context has often been limited due to perceptual channels and aesthetics, to the point where most point estimation is not possible. In addition, the active component of graphs is often weak in

auditory graphs, where a playback of a sound file is the common practice.

However, given these limitations, auditory graphs still show much promise. Building on the early work of others [? ?], and expanding auditory graphs to provide more interaction and possibilities for point estimation, auditory graphs could be used for many real-world accessibility situations in the classroom. This thesis will demonstrate that:

Proposed Thesis Active point estimation with sonification enables people to solve middle school graphing problems, regardless of visual impairment.

“Active point estimation” (APE) is the process of guessing the value of a piece of data, in a user-directed fashion. In visual graphs, active point estimation is facilitated by the eyes, which flick between the tick marks and labels on the axis and the data point being explored⁵. In other formats, the perception may be different, and may additionally require feedback, such as mouse movement. However, the concept of using the body to explore the graph remains the same. Importantly, the conversion of non-verbal information into a verbal mapping is a critical part of this process. Obviously, if the information could be fully communicated with text or speech, that should be the sole way the information is delivered, as it is more concise and easier to create. Most graphs do have some verbal components, used to label important components of the context and plotted data; these components can be available in other formats. Active Point Estimation with Sonification (APE-S) is the use of sonification as the format for displaying the non-verbal information. Speech can be used with APE-S for the verbal parts of the data display.

“Middle school graphing problems” are graphing questions that are based on middle school mathematics standards. For this research, the standards used are the CCS for Mathematics and the GPS, particularly those standards found in grade 6. This grade is a good starting point for several reasons. First, it is at a point where there are many one dimensional number line problems and two-dimensional graphing problems, so the graphing approach can be evaluated in these two important numbers of dimensions. Second, sixth grade is in the middle of “K-12”, so it will be a good test for the efficacy of the approach.

⁵This is intuitive, but cite something about this.

If it works, then it will appear that about half of the K-12 standards for graphing work. If something does not work, it can be evaluated at a lower grade level at a later time.

These graphing problems are also limited by the way they are supposed to be solved. Specifically, students are required to solve them by hand, with the use of paper, worksheets, or graphing paper. Students must learn how to use the graph basics before depending on a calculator for more advanced graphing functions. This provides an opportunity to evaluate technology as an assistant for basic graph reading and writing, instead of as a crutch for learners who cannot see (or cannot calculate).

Finally, active auditory graphs can be used by people regardless of their visual impairment. Students who are blind, and understand the mathematics, are able to solve graphing problems with the tool. In addition, students who are low vision or sighted can also use the tool to solve graphing problems. Importantly, for all of these students, the tool facilitates graphing to the extent that graph paper helps. The tool is designed for a student to read and write a graph, not to replace a student's graph literacy.

This thesis covers three important areas. First, the use of sonification for active point estimation has open questions for input and output devices. Second, the components required for a sixth grade graphing tool are unknown; based on standards and a task analysis, a set of requirements will be developed. Third, an ecological evaluation of the resulting system supporting the required components showed the efficacy of the solution. These are the three phases of the upcoming chapters, and form the body of the research for this dissertation.

The research questions for the phases are in Table 1. Questions R_1 and R_2 focus on the input and output devices available for active point estimation. Important questions include the use of the mouse or keyboard as input, potential target sizes, and the way the sonification should be displayed for maximum targeting speed and accuracy.

The next four research questions explore the standards (R_3), graphing questions (R_4), steps (R_5), and potential alternative presentations (R_6) of graphs. This novel process, called SQUARE, provides a way to create alternative technologies that can retain the core concepts and learning found in the original approach.

The final questions look at the new tool in a classroom setting, considering setup (R_7),

Table 1: Research Questions

| | |
|-------|---|
| R_1 | How can auditory display facilitate interactive point estimation? |
| R_2 | What common input devices can be used by blind people for interactive point estimation? |
| R_3 | What education standards require graphing? |
| R_4 | What are example graphing problems that meet each standard? |
| R_5 | What steps are used to solve the graphing problems? |
| R_6 | How can an accessible auditory graphs tool enable the steps necessary to solve the graphing problems? |
| R_7 | What issues are there in preparing classroom materials with an accessible auditory graphs tool? |
| R_8 | What issues are there in using an accessible auditory graphs tool in classroom situations? |
| R_9 | What issues are there in using an accessible auditory graphs tool in testing situations? |

classroom issues (R_8), and test results (R_9). It includes results from both low vision and blind participants.

Through this dissertation, I will report on auditory graph compatibility with course graphing tasks, interfaces used to meet the task goals, and measurements of student performance. I will also have an auditory graphing tool, GNIE, that visually impaired students can use with coordinate graphs.

This dissertation has several contributions. This project will identify whether auditory graphs can be used for point estimation. This information can be combined with the large body of research in trend analysis and sonification to build a complete graphing system.

The SQUARE process itself is a contribution. It provides a way for researchers to develop other assistive technologies, or expand the math technology to other grades and mathematics standards. The results of the SQUARE method are a further contribution. They will build the basis for a bottom-up understanding of graph literacy in K-12 education in the United States. In the long term, this can be used to understand the development of graph literacy and further tune graph learning. For example, the results show that there are a handful of actions a student must know to complete all sixth grade graphing problems. A teacher can build graph literacy by ensuring his students know the basic actions.

The combination of the active point estimation understanding, along with the results of SQUARE, led to the development of GNIE, an assistive technology for graphing.

Table 2: Contributions

- | | |
|-------|---|
| C_1 | A greater understanding of how interaction devices and auditory display choices affect speed and accuracy for point estimation. |
| C_2 | SQUARE, a method for creating education technologies based on graphing standards. |
| C_3 | A list of standards, questions, and steps based on requirements for sixth grade mathematics. |
| C_4 | GNIE, an auditory graph technology that can be used to solve graphing questions from many Common Core Standards. |
| C_5 | High-Low, a method for evaluating assistive educational technologies as testing accommodations. |
| C_6 | Observations and feedback about preparing and using auditory graphs in classroom settings. |

Genie! (Genie!) is a contribution in itself, as the tool has been proven to be in line with sixth grade common core graphing standards.

The CCS will be used to produce a list of graphing standards for high school mathematics students. It will also present steps to solve questions based on the standards, and an auditory graphing tool that enables students to complete those steps. The first part provides a way for evaluating whether a graphing technology is suitable for a high school mathematics classroom. The summative evaluations will produce evidence of effective methods for evaluating auditory graphs in a classroom setting and evidence of effectiveness of auditory graphs on a test, compared with other formats.

The proof of concept occurred during Phase 3, evaluation. The first major contribution is High-Low, a method for evaluating whether assistive technologies are suitable testing accommodations in education. Second, observations and feedback about how to prepare and use auditory graphs in classroom situations will be a valuable resource to researchers, educators, and students.

1.3 Document Overview

This introduction gave a brief overview of graphs, graphicacy, mathematics education, graphs for the blind, and auditory graphs. It also described the thesis, research questions, and contributions of this dissertation. Chapter 2 describes related work in more detail. Chapter 3 describes the technology. Chapters 4 to 6 discuss the three phases of the

research. Chapter 7 is general discussion, and chapter 8 concludes the thesis.

CHAPTER II

BACKGROUND

This dissertation explores how to improve learning opportunities for visually impaired students using graphs. Graph literacy is the ability to understand and create a coordinate graph or number line. This literacy, dubbed “graphicacy”, has been suggested as a key component in mathematics education, in line with reading, writing, and arithmetic [?]. A look at learning standards shows graphs and number lines used at every grade between Kindergarten and 12th grade (figure 3 in chapter 1) Graph literacy is important beyond school as well, as a critical skill in many white collar jobs. With such a high demand for graph literacy, it should be clear what constitutes graph literacy and the development of a student’s graphicacy through their K-12 education.

However, there is surprisingly little documentation about the component parts of graph literacy. Education standards such as the current Georgia Performance Standards (GPS) and the upcoming Common Core Standards (CCS) mention graphs in many places, and curriculum is peppered with graphs and number lines. However, there is not a clear structure of graph literacy development, nor a suggested progression of student learning of the pieces over the K-12 education years. Graphicacy theoreticians are quick to point out the importance of graphs with insightful examples [? ?], but their explanations are not clearly tied to everyday graph interpretation in K-12 classrooms. Several researchers have created important works in accessible graphs and charts [? ? ? ? ? ? ? ? ?]. While many evaluate their systems, they are not clearly tied to activities that students do in a classroom or a test. And while auditory graphs have had a slow, 30-year development, the problem of point estimation remains in most prototypes. This chapter suggests a comprehensive analysis of what graphicacy means for students has been lacking; a deeper understanding of “graphicacy” would serve as bond, combining classroom practice, graphicacy theory, auditory graph perception, and assistive technology development. Defining what graph literacy

is, in terms of actual mathematics standards and actual steps to solving graphing problems, will lead to an understanding of the fundamentals of graphicacy, which in turn can be used to design relevant classroom curriculum and assistive technologies.

This chapter has three major sections. The first section introduces auditory graphs, and explores their application for education. While the interactive components of auditory graphs have been discussed (e.g. [?]), they have not been widely used for point estimation. While there have been some promising first steps, previously proposed technologies have not been evaluated to be in line with education standards or curriculum. The second section describes education curriculum, graph literacy, testing accommodations and the current practice of how visually impaired students learn graphs. Tyler's four steps [?] outline the process for developing and advancing any curriculum. This will be adapted to explore the integration of assistive technology. The final discussion section integrates the key points of the related work in order to move forward with a comprehensive plan for discovering the tasks involved with graph literacy.

2.1 Graph Literacy

Graph literacy, or graphicacy, is an important part of mathematics education [? ?]. As seen in Figure 3 (Chapter 1), graphs compose 5-25% of the GPS every school year from Kindergarten to grade 12. This section explores graph literacy. It suggests why graphs are different than algebra, in practical and theoretical terms. While graph literacy proponents have developed useful examples for their cause, they have not sufficiently explained graph literacy in terms of everyday use, or the functional components of graph literacy. This section begins with a demonstration of how graphs can give insight that is not apparent in data or statistics.

2.1.1 Anscombe's Quartet

In 1973, Anscombe [?] presented four data sets of x,y pairs with 11 data points, where visual analysis of the written table had no obvious differences (see Table 3). Several descriptive statistics are the same between the four data sets [?]:

Table 3: Anscombe's quartet, data table. See figure 5 for the graph of this data.

| Obs. no. | Data Set 1 | | Data Set 2 | | Data Set 3 | | Data Set 4 | |
|----------|------------|-------|------------|------|------------|-------|------------|------|
| | x | y | x | y | x | y | x | y |
| 1 | 10 | 8.04 | 10 | 9.14 | 10 | 7.46 | 8 | 6.58 |
| 2 | 8 | 6.95 | 8 | 8.14 | 8 | 6.77 | 8 | 5.76 |
| 3 | 13 | 7.58 | 13 | 8.74 | 13 | 12.74 | 8 | 7.71 |
| 4 | 9 | 8.81 | 9 | 8.77 | 9 | 7.11 | 8 | 8.84 |
| 5 | 11 | 8.33 | 11 | 9.26 | 11 | 7.81 | 8 | 8.47 |
| 6 | 14 | 9.96 | 14 | 8.1 | 14 | 8.84 | 8 | 7.04 |
| 7 | 6 | 7.24 | 6 | 6.13 | 6 | 6.08 | 8 | 5.25 |
| 8 | 4 | 4.26 | 4 | 3.1 | 4 | 5.39 | 19 | 12.5 |
| 9 | 12 | 10.84 | 12 | 9.13 | 12 | 8.15 | 8 | 5.56 |
| 10 | 7 | 4.82 | 7 | 7.26 | 7 | 6.42 | 8 | 7.91 |
| 11 | 5 | 5.68 | 5 | 4.74 | 5 | 5.73 | 8 | 6.89 |

Each of the four data sets yields the same standard output from a typical regression program , namely

- Number of observations(n) = 11
- Mean of the x 's (\bar{x}) = 9.0
- Mean of the y 's (\bar{y}) = 7.5
- Regression coefficient (b_1) of y on x = 0.5
- Equation of regression line: $y = 3 + 0.5x$
- Sum of squares of $x - \bar{x} = 110.0$
- Regression sum of squares = 27.50 (1 d.f.)
- Residual sum of squares of $y = 13.75$ (9 d.f.)
- Estimated standard error of $b_i = 0.118$
- Multiple $R^2 = 0.667$

However, graphs of each of the data sets show great differences. As seen in figure 5, the graphs show different levels of correlation to different types of fit lines and curves. The trends in the graph are intuitive, and, once seen, can be described. A text summary can be created. However, before looking at the graph, our common data analysis and statistics did not find differences, so the text description could only be created *after* looking at the

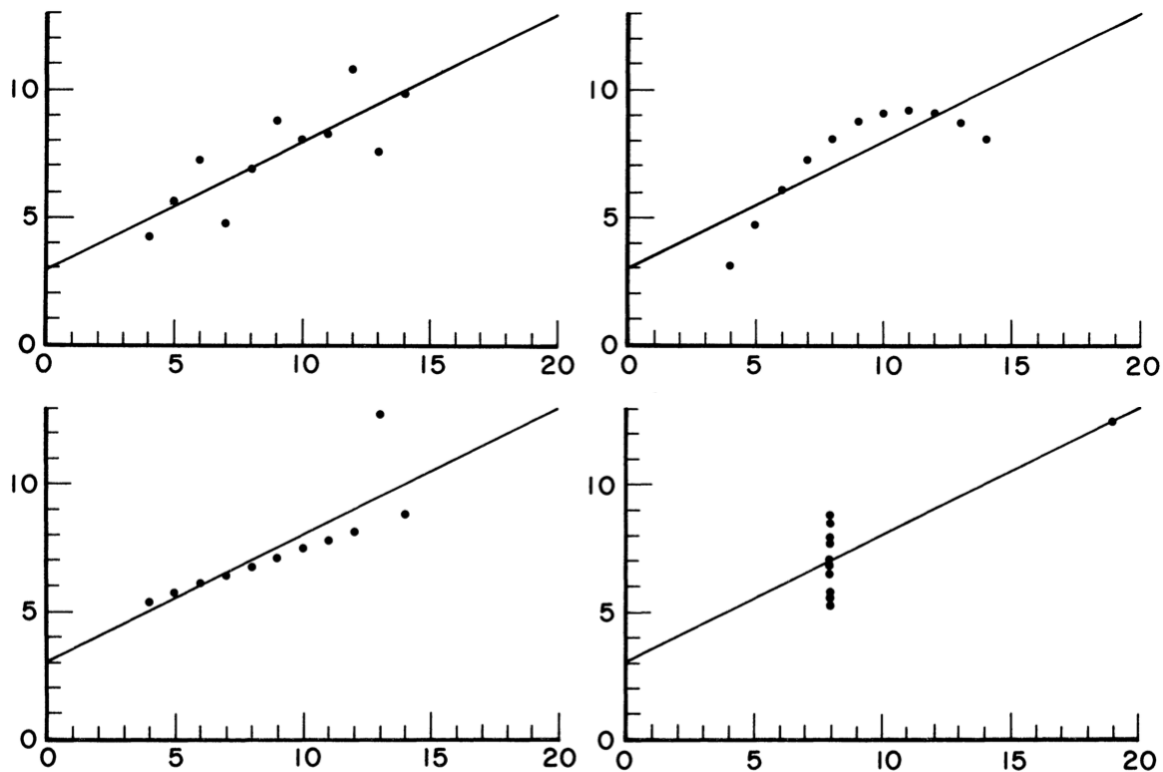


Figure 5: Anscombe’s quartet of visual graphs [?]. The original data is in table 3. All of these graphs have the same x and y mean, variance, correlation, and line of fit. However, the graphs are visually distinct, with different types of trends. The top left shows a moderate linear correlation between x and y, slowly increasing. The top right shows a parabola-like curve facing downward. The bottom left is a strong linear correlation between x and y, with a single outlier. The bottom right is a vertical line at $x=7$, with an outlier.

graph. Anscombe’s Quartet [?] shows how graphs can give insight that is not available by looking at the original data or using statistics ¹. In other cases, information may be available in both formats, but it may be more efficient to use graphs over other methods.

2.1.2 Graphs and Efficiency

Even if the information is available in both formats, a graph may be more time efficient than using a verbal description or formula to solve a problem. Larkin and Simon explain why diagrams (including graphs) take less time than reading a data table:

When to two representations are informationally equivalent, their computational efficiency depends on the information-processing operators that act on them.

¹Data with properties like Anscombe’s Quartet but with different raw values can be generated. See [?] for an algorithm.

Two sets of operators may differ in their capabilities for recognizing patterns, in the inferences they can carry out directly, and in their control strategies (in particular, the control of search). Diagrammic and sentential [verbal] representations support operators that different in all of these respects. Operators working in one representation may recognize features readily or make inferences directly that are difficult to realize in the other representation. Most important, however, are differences in the efficiency of search for information and in the explicitness of the information. In the representations we call diagrammatic, information is organized by location, and often much of the information needed to make an inference is present and explicit at a single location. [...] Therefore problem solving can proceed through a smooth traversal of the diagram, and may require very little search or computation of elements that had been implicit.

Larkin and Simon [?] explain why graphs are faster, but over-emphasize location. More generically, data is converted to a different format through the process of *mapping*. For visual graphs, location is often used as a mapping. There are other common mappings, such as color, size, and shape. In tactile graphs, common mappings are location, texture, 2-D shape, height, and size. In auditory graphs, common mappings are pitch, timbre, pan, and volume (see section 2.2 for more details). Like location, these mappings can make it faster to find insight than by simply looking at a data table.

Location as a mapping is not strictly only about spatial location. Graph readers understand that there are unstated rules that bend a strict interpretation of data-to-location. There are many examples. The size of a point takes up space: the point represents a single value, but, if the space is taken literally, it encompasses a range of values within the circle-like region. Tick marks show their location, but their actual indicator is at the intersection of the tick mark line and the axis line. Labels are offset from their target location (often grouped mentally by a proximity Gestalt), so that someone may read both the label and the point. While it is important to limit the definition of a “graph”, it also must be flexible enough to encompass alternate formats that are conducting the same function, specifically tactile graphs and auditory graphs. In particular, it is not necessary for an auditory graph

to be spatial. It is only necessary for the auditory graph to have mappings that enable similar tasks, speed, and types of insight.

While previous work in graph literacy has created insightful reasons to use graphs, the link between day-to-day graphs use and theoretical graphs use has not been sufficiently explored. Such a study will lead to two contributions. First, the theoretical contributions of graphs can be tested in real environments. Do people use graphs for trends? Is it time efficient? Second, the practical use of graphs can inform the theoretical contributions. Are there ways people use graphs that have not been sufficiently explained theoretically? What are the building blocks of graph literacy?

In addition, since this dissertation is focused on alternative formats, the theoretical guidelines for graph literacy can be used to evaluate a new technology. For example, students should be able to use the new technology much like existing technologies, in terms of the steps needed to complete the graphing problems, and the efficiency of the completion.

This discussion of graph literacy has introduced two alternative formats: tactile graphics and auditory graphs. Auditory graphs are not often used in education, and will be discussed next.

2.1.3 Discussion

Graphs provide an alternate, non-verbal format for viewing numerical data relationships, while maintaining sufficient verbal information to estimate the original values. In auditory graphs research, the most effective data mappings are non-spatial. Any alternate format, however, should consider the learning goals of graphing, and should not replace an opportunity to gain graph literacy.

2.2 Auditory Graph Fundamentals

The proposed technology is auditory graphing software for visually impaired students. To begin, an understanding of auditory graph basics is necessary.

There are a few terms related to the concept of “auditory graph”. “Sonification” is the use of non-speech audio to convey information. If the data itself is directly played as an audible sound, such as a waveform, then the sonification is an “audification,” [? ?].

In many cases, however, there is a mapping of the data into an auditory form that will display more perceptible differences in the data; this process is called “parameter mapping” [?]. An “auditory graph” is a sonification with parameter mapping that is “the auditory equivalent of mapping data to visual plots, graphs and charts,” [?]².

Like visualization, properties of the auditory medium can be manipulated in ways that people can easily perceive. In visual graphs, visual properties such as spatial location, color, size and pattern are often modified to convey information. For a sonification, audio properties such as pitch, pan, rate, volume, and timbre may be modified. Also like visualizations such as coordinate graphs, sonification can have verbal (spoken) components, yet the non-verbal components are a critical part of the display.

2.2.1 Acoustics and Psychoacoustics

Acoustics is the study of the mechanical movement waveforms traveling through particles in materials. The largest and most relevant component of acoustics is the study of how sound is produced, propagates, and physical properties of the sound waveform.

Psychoacoustics is the study of the perception of sound. The sensation and low-level interpretation of sounds results in an understanding of the sound in ways that are different than what was actually produced. Psychoacoustics, to some extent, also explores how people make meaning of the perceived sounds.

2.2.1.1 *Anatomy*

The human ear has three major parts [?]. The outer ear is composed of the pinna (the visible “ear”) and ear canal. The middle ear has several ear bones which change the amplitude of the incoming waveforms, and transmits the new waveforms into the inner ear. The inner ear is filled with fluid, which transmits the waveforms. Hairs on the basilar membrane move when a wave passes over them; this movement triggers neurons to fire and transmit to the brain. The basilar membrane and the receiving neurons in the brain have

²Vickers [?] further distinguishes auditory graphs from sonified graphs, with respect to their level of abstraction from the data. This distinction will not be particularly relevant for my research. For the purposes of this dissertation, the only term used will be “auditory graphs”.

a tonotopic mapping, meaning certain tones are located at particular physical locations of the cochlea and brain [?].

2.2.1.2 Acoustic Properties

Researchers often create sounds in order to study them. One of the simplest sounds, in terms of controlling the properties of the waveform, is a sine wave. With a one audio speaker, a generated sine wave has a certain frequency and amplitude. Frequency is the number of compression waves the sound has at a particular listening point over the course of one second [?], also known as Hertz (Hz).

Amplitude is the amount of change in pressure of the wave, seen visually as the height of the sine wave. Amplitude is often represented in *dynes/cm²*, or in decibels (dB), a logarithmic scale [?].

A third property in acoustics is the complexity of the sound [?]. Natural noises do not sound like a sine wave. They often can be characterized with several frequencies with varying amplitude, and changing waveforms over time. Several interesting properties of complex sounds exist, such as hearing only parts of the waveform, noise cancellation, hearing beats when presented with two similar frequencies, and masking [? ?]. These are important properties of sounds, but will be avoided for the graphing display by using redundant cues, complex MIDI notes, and separate timbres.

2.2.1.3 Perception

Pitch is the perception of frequency (mostly) [?]. Measured in mels, pitch changes in a complex way based on frequency, and to some extent amplitude and tone complexity. Stephens et al. [?] asked participants to use a knob to change frequency, and indicate when the perception of the frequency was cut in half. They found a relationship to frequency summarized in Figure 6. For practical purposes, pitch can be thought of as musical notes, scaled to double every 12 semitones (1 octave). A mapping of musical frequency can be simplified as

$$f = 2^{N/12} * 220Hz \tag{1}$$

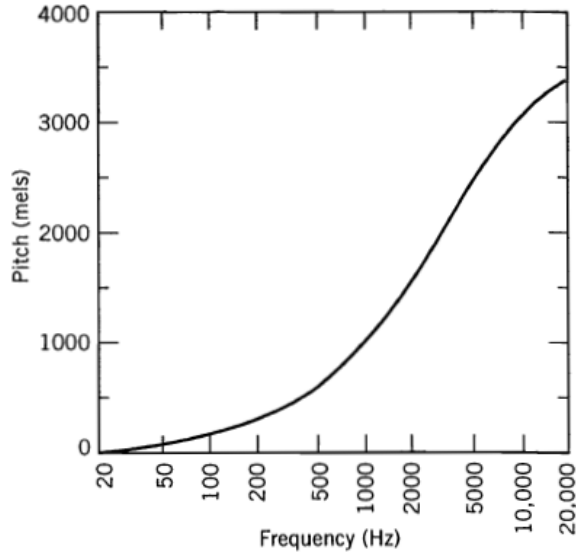


Figure 6: The relationship between frequency and pitch, as measured by Stephens et al. [?]. As noted in [?], the pitch changes quickly for low frequencies and slower for high frequencies. Note the frequency scale is logarithmic.

N is the semitone-difference from A3 (for example, B3 is +2, since it is two semitones from A). A3 is defined as 220 Hz. For each octave, frequency doubles [?]. Since there are 12 semitones per octave, N is divided by 12 in the exponent of base 2. f is the resulting frequency.

Loudness is the perception of amplitude. One sone is the loudness of a 1000 Hz sine tone at 40 dB SPL. When intensity triples (increases 10dB), the loudness doubles. Loudness is also greatly affected by the tone’s frequency [?]. People are more sensitive to sounds between about 200-5000 Hz, and will perceive sounds with the same amplitude but outside of this frequency range as quieter.

Localization is the perception of the position and distance of sound sources. Localization is supported with many aural cues, and works best with two functioning ears. For the purposes of sound design in this dissertation, localization will be limited to left-right stereo panning in headphones and speakers.

Timbre is the characterization of sound complexity into types of sounds. As noted by Schiffman, “the fundamental frequency mainly determines the *pitch* of a complex sound, whereas its harmonics determine its *timbre*,” [?] (original emphasis). Musical instruments

differ by their timbre.

2.2.2 Mapping

A mapping is “the dimension of sound that is employed to vary with and thus represent changes in data,” [?]. Since 1985 [?], most auditory graphs have used a pitch mapping for y-axis values (higher data is higher pitch) and a time mapping for x-axis values (higher data is later in time)³. It is easy for many people to perceive small changes in pitch, and cognitively map these changes to the data, in terms of understanding the general slope of the line or correlation to a scatterplot [? ? ? ?]. For interactive graphs, or in addition to the time mapping, pan is sometimes used to map to x-data values (far left is lowest data, far right is highest data). When there are more than two dimensions, designers often produce sonifications that map to other dimensions, such as rate and volume. The efficacy of these sonifications and the use of other sound types in general, however, is more difficult to understand to the user. Based on a career in auditory interfaces and cognitive psychology, Flowers [?] states:

Listening to simultaneously plotted multiple continuous pitch mapped data streams, even when attention is given to timbre choice for different variables to reduce unwanted grouping, is probably not productive. It is possible that with levels of consistent practice that are well beyond those of most sonification evaluation studies, we might do somewhat better at listening to multiple sonified streams than is currently apparent. But it is generally the case that attending to three or more continuous streams of sonified data is extremely difficult even when care is given to selection of perceptually distinct timbres or temporal patterning.

It appears that auditory design must be conducted carefully, so that the bandwidth and quality of information presented is manageable by most users. While Flowers appeared to be discussing the actual data values, other non-speech sounds such as tick mark locations

³The direction that lower or higher data maps to properties of the auditory display is called polarity. For more on polarity and its applications, see [? ? ? ?].

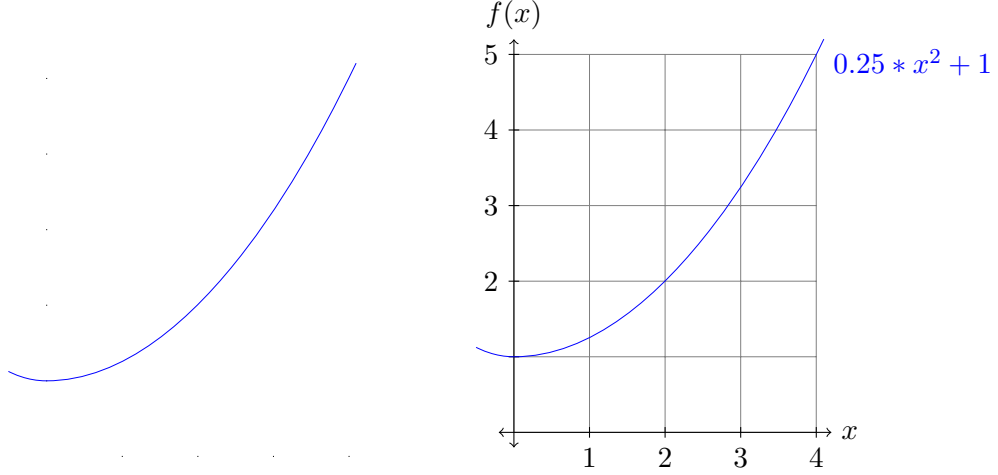


Figure 7: The graph of $f(x) = 0.25x^2$, on the domain of $-0.5 \leq x \leq 4.1$ with and without context. In both cases, it is possible to see where the graph is increasing, decreasing, and flat. The function also looks like a parabola. However, the first image cannot be used to determine intercept or any of the points.

probably also make it more difficult to perceive the changes for any of the active streams.

One important aspect of mapping is the mapping process. In a typical mapping, a group of data is converted to psychoacoustic properties in a linear fashion. For example, a range of data values from 5-213 could be mapped to a range of MIDI notes, such as the recommended range 35-100 [?]. With a positive polarity, the data value 5 would be represented by MIDI note 35, and 213 would be represented by MIDI note 100. A linear scaling would occur in the middle. More generally, the mapping would be (for positive polarity):

$$N_{value} = \lfloor (((D_{value} - D_{min}) * (D_{max} - D_{min})) * (N_{max} - N_{min})) + N_{min} \rfloor \quad (2)$$

In words, the location of the data point within the range of all the data, as a fraction from 0 to 1, is mapped onto the range of note data, in a linear fashion.

2.2.3 Context

At first glance, the data-to-sounds mapping in the previous paragraph may appear sufficient to interpret the graph. A listener could, for example, hear if the data values were increasing, or determine the graph family. For most practical situations, however, more information

is necessary. It is impossible to find the intercept or any other points if only shown the data. Consider Figure 7. The two graphs appear identical in form. However, the first is actually a possibility for an infinite number of upward-facing parabolas. Similarly, if bound to constant visual ranges, a set of increasing values will sound like it is increasing, but point estimation is simply guessing. $y = x$ sounds like $y = 2 * x + 3$. In plotting the point, critical information about the data was lost.

The missing component is context. Context is the presented relationship between the line and the data values. In Figure 7, there are several pieces of context: each axis (x and $f(x)$) had an axis line, tick marks, grid lines, tick mark labels, and an axis label. The line also had a label. Non-verbal components such as the axis, tick marks, and grid lines assist in finding relating a specific point to the spatial range. The tick mark labels relate the tick marks and grid lines to data values. Therefore, a sighted person can look at where the point crosses the a grid line, follow the grid line to the label, learn one of the two data points for that piece of the graph line, and repeat with the other dimension. The viewer can also interpolate the spatial distance to represent a “data” distance⁴, so that when x crosses 3, y is at about 4.2.

In many cases, there is a natural mapping of magnitude estimation based on the type of data involved. In a series of reports, Walker and others [? ? ?] asked participants to gauge the magnitude and direction of the sounds, given particular types of data (such as temperature or dollars). They found significant differences in the slope and polarity for the different data dimensions. Thus, with knowledge about the data, sonifications can be optimized for the easiest interpretation⁵. However, this natural mapping is not sufficient for detailed point estimation.

Auditory equivalents to tick marks are possible. Smith and Walker [?] used x-axis percussion clicks and various implementations of y-axis tones to determine which designs led to the highest reduction in errors. The graph represented the price (dollars) of a stock

⁴Arithmetic of course gives the exact answer, $0.25 * (3)^2 + 1 = 3.25$. But data tables have their own downsides, such as Anscombe’s Quartet, described in Section 2.1.

⁵This also depends on the user group. Walker and Lane report different polarities with visually impaired participants, when compared with sighted participants [?].

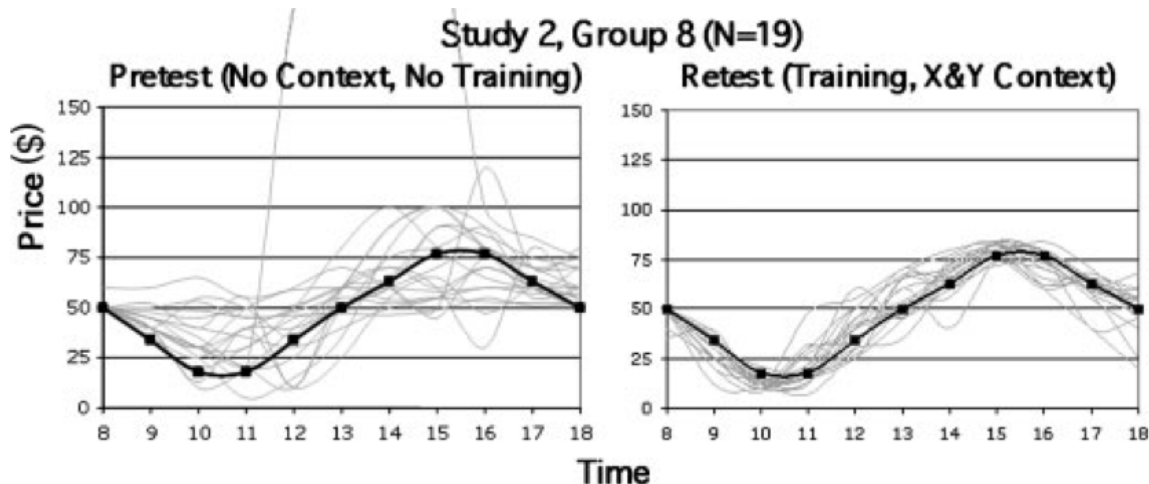


Figure 8: A portion of figure 4 from Smith and Walker [?]. Note that the point estimation of each participant's recreation of the black line (in gray) is close, but probably varies from the correct answer more than what a mathematics teacher would judge to be correct.

over a 10-hour trading day. The x-axis data was mapped to time and the y-axis was mapped to pitch. In some conditions, the x-axis had a click on each hour, with one click per second. For each condition, the y-axis context had one of the following: a constant tone representing the opening price of the stock; a dynamic reference tone at the day's high (\$84) and low (\$10), playing the low tone when the price was falling and the high tone when the price was rising; or no context. Findings showed a significant reduction in the number of errors for the dynamic conditions, indicating that the context was helpful in estimating points. Follow up studies have shown that context can be improved with training [? ?].

The series of studies by Smith and Walker [? ? ?] provide solid research at the use of context in auditory graphs. In terms of practical situations, however, the error rate is simply too high. Consider Figure 8, from [?]. These results show a visible decrease in variance of the drawing from the pretest to the retest. However, in terms of math-class correctness, the lines in general do not appear to be sufficiently correct. In [?], the mean absolute error for each trial was \$6.420, or about 8.6% of the range of the graph. While better than the control group (\$11.798, or about double), this is the error for *each trial*; in a classroom setting, students will have to read and write every point on a graph⁶.

⁶What would teachers judge an acceptable level of error for understanding graphs? In Phase 3, studies 2 and 3 (starting in Section ??), I propose that peoples' accuracy with auditory graphs can be compared with

The clearest way to provide more context is with speech⁷, but speech has its own drawbacks. Rather than relying on what is essentially an earcon (a non-speech, arbitrary⁸ representation of information [? ?]), which is then mapped to a verbal meaning such as “10 dollars”, speech can be used to indicate when someone is at a dollar mark. Unfortunately, speech makes it more difficult to use a fixed pitch, which was the critical component to the reference earcon-tones⁹. In addition, since speech can take a relatively long time for listening, the x-axis timing element can become complicated. In Smith and Walker’s first study [?], there was one second between tick marks. It may be ambiguous where on the graph a spoken “eighty-four” relates to the sound stream being produced.

One possibility is the use of interaction in auditory graphs. Instead of having a playback, have the user control the x and/or y-axis, and play the sounds representing the data at that point. For example, a pitch could be played representing the y-axis, while the x-axis is mapped to the mouse position instead of time. Verbal cues, then, can probably be understood to relate to a certain point by simply slowing down the movement around where a statement is spoken, until it is clear to the user what the point represents. Since there are two axes, perhaps there are two voices.

One aspect of this change is that the x-axis is no longer clearly mapped to anything non-visual, other than a relative horizontal movement of the mouse. Perhaps a second pitch could be played alongside the y-axis pitch, or perhaps the data could be combined, such as x-axis rate and y-axis pitch. It could also be the case that the verbal information is sufficient.

One important final note on the use of context is the use of a window. A graphing “window” is the data context range shown for each dimension. Many formulas cannot be entirely plotted, since they would take a large or infinite amount of space. Instead, the

their accuracy in tactile and visual graphs.

⁷Almost every visual graph has verbal (text) components. Since the Smith and Walker studies were concerned with the use of non-speech audio for context, speech is rightfully excluded. But for practical purposes, it should be an part of auditory graphs.

⁸In this case, the notes are not completely arbitrary, since their values actually hold information, such as pitch-to-data. But the representation of, for example, “10 dollars” to the musical A3 must be remembered.

⁹Text-to-speech engines can be manipulated to produce a narrow pitch range and a fixed pitch of sounds, but monotone voices are harder to understand, and the perception of the pitch and its relationship to other tones may remain difficult.

graph’s author selects a manageable range where the important aspects of the graph are perceptible. For example, in Figure 7, the x-axis shows the range 0 to 4 and the y-axis shows the range 0 to 5. In other cases, the same graph may be displayed with a different window. It is also simply not enough to take the minimum and maximum bounds of the graph. For example, when comparing two graphs, side by side, it is useful to have the same window for each graph, so that they can visually be compared. Otherwise, for example with a min-max window range on the y and a fixed window range of 0 to 10 for x, the functions $y = x$, $y = 10 * x$, and $y = 0.01 * x - 20$ all have a sloped line going from the bottom left to the top right.

2.2.4 Trend Analysis

Trend analysis is the process of understanding the shape and direction of data. The earliest work in auditory graphs demonstrated that people could perceive the mapped trends. In 1985, Mansur and Blattner [?] reported that participants could identify graph symmetry, monotonicity¹⁰, and approximate slope of auditory graphs.

2.2.5 Point Estimation

Point estimation is identifying the numerical (or categorical) value of a point. Trend analysis and point estimation can be thought of as two parts of graphing; taken together, they appear to comprise the whole of graph perception. With Sound Graphs [?], it became clear that trend analysis would work with auditory graphs, and 30 years of follow up studies have shown a wide range of possibilities. However, point estimation remains a challenge. It is simply difficult to understand where a number is located. In fact, point estimation may have a wider scope in graphs than trend analysis. Many graphing problems in education, such as “plot (2,3),” require point estimation with no trend component¹¹.

¹⁰ A monotonic function always increases or always decreases, but the slope at different points can change.

¹¹ While technically true, there is probably a level of trend detection in place. For example, it is easy to tell with a visual or auditory graph that has axis context that the tick mark is to the right of the y axis. However, this trend is a useful check on solving the problem correctly, and not fundamentally necessary for answering the question.

2.2.5.1 Fitts's Law and finding targets

In HCI, the speed to finding a target is often predicted by Fitts's Law. Fitts' Law is a movement time prediction, based on target distance and size, and device-specific constants. The particular formula varies depending on the context, but is essentially:

$$MT = a + b * \log_2 \left(1 + \frac{D}{W} \right) \quad (3)$$

where MT is movement time, D is distance to target, W is width of target, and a and b are device-specific constants obtained in empirical studies [?]. For a given device, the time it takes to get to a target varies directly on the distance: a longer distance leads to a longer time. The movement time varies inversely to the target width: a wider target leads to a shorter time. The original Fitts' Law study used a physical pen to move between targets [?], and subsequent studies have shown it is applicable for screen movement with the mouse and other devices [? ? ?], in one and two dimensions [? ?], and for accessible interfaces [?].

2.2.6 Conclusion

The cognitive aspects of auditory graphs are bound by the perceptual system. Since the 1980's, designers have mapped x-axis data in the basic auditory graph to time and y-axis data to pitch. In lab studies, visually impaired and sighted students can detect trends and estimate points with sonifications. Some aspects, however, may have unacceptably low accuracies for practical purposes. Point estimation is particularly difficult, so alternative forms of point estimation, including interactive control and speech feedback, should be explored before system development.

2.3 Auditory Graphs as Educational Technology

Auditory and multimodal graphs have been explored as tools for education in a number of research projects. Each of the five projects discussed add a useful background to developing a practical system. However, there is insufficient ties to curriculum, so it is impossible to determine the practical utility of the proposed system in K-12 education.

2.3.1 Sound Graphs: Successful Trend Analysis

The published interest in auditory graphs for visually impaired dates back to 1985. “Sound Graphs” [?] presents the problems with the status quo tactile graphics, defines what is now the typical auditory graph, and gives some evaluation results¹². Mansur and Blattner saw the goal of the work to provide people who are “blind with a means of understanding line graphs in the holistic manner used by those with sight,” [?]. As stated, the research focused on holistic types of evaluation, including line slope, graph family (lines or exponentials), monotonicity, convergence, and symmetry. They found that in speed and accuracy, participants were nearly equivalent to tactile graphics (83.4% for audio and 88.3% for accuracy, significantly different). However, Mansur and Blattner considered auditory graphs to be superior in production, as tactile graphics are “slow and difficult to use, require considerable time to engrave, and are fairly inconsistent in quality,” [?]. While the paper described a simple playback approach, they suggested more interactive methods: “the blind user must be capable of controlling the system via keyboard, joystick, “mouse,” or other means,” [?].

Mansur and Blattner introduce a few themes common in today’s accessible graphing research. Many authors highlight the use of auditory graphs for visually impaired users (e.g. [? ? ? ?], but some specifically work with sighted students [?]). Tactile graphics, still the status quo, remain the key comparison in modern studies. The authors also propose several additions of sound, such as “the points where the global maxima and minima occur [. . .] inflection points, discontinuities in the curve, or the point where the curve crosses some y-value,” [?]. These support cues are similar to the y-axis minimum and maximum context evaluated 20 years later in Smith and Walker [? ?] (Section 2.2.3).

The selection of slope, family, monotonicity, convergence, and symmetry appears to not be empirically based on graphs in education, or graphs in the workplace. In fact, the application of certain graphing components in several accessible graphing projects appears to be largely unguided by what is important in these domains.

¹²While Bly [?] first introduced auditory graphs, Mansur and Blattner specifically approach auditory graphs as assistive technology.

Mansur and Blattner provided the first exploration into the accessible auditory graphs space. Further research in the area became more ecological and wider in scope and technology.

2.3.2 Plotting points with the Integrated Communication to Draw

Hesham Kamel, a blind PhD student at Berkeley, explored “computer-aided drawing for the visually impaired,” [?]. This work led to a multimodal system, the Integrated Communication 2 Draw (IC2D) that was the core of Kamel’s thesis and early publications (with James Landay and others) [? ? ? ? ? ? ?].

Early work on the IC2D outlined its benefits. In terms of input,

The IC2D provides access, using the computer keyboard, to nine fixed screen regions in a 3x3 grid corresponding to the numbers on the keypad [...and] a recursive scheme to provide the user with a hierarchy of grids, allowing a more refined resolution of navigational access. [?]

In the IC2D, a blind person could use the keyboard to find a point “analogous to pointing and clicking with a mouse for a sighted user,” and intentionally leave that point and easily find it again [?]. The interface used both relative and absolute positioning options with certain key commands. Speech was used to indicate positions, such as “position 7, bottom right, middle left” (for a recursion depth of 3) [?]. In addition, since adding a point involved using a series of keyboard commands, the commands could be easily repeated in other parts. For example, drawing a line in the top three cells of the upper left square would involve pressing 1-m-1-k-2-k-3-k-(info)-(line), and repeating the commands in cell 2 (extending the line) would involve pressing 2-m-1-k-2-k-3-k-(info)-(line), so similar output is based on similar input. It is very rare for blind people to have access to such graphics tools; the capability of the IC2D to allow visually impaired people to read and even plot on a canvas was groundbreaking. Later work showed how the IC2D grid concept could be used with earcons [?] and even animation tools [?]. Figure 9 provides a visual example of IC2D.

Follow-up research uncovered specific goals for accessible drawing tools. Based on a

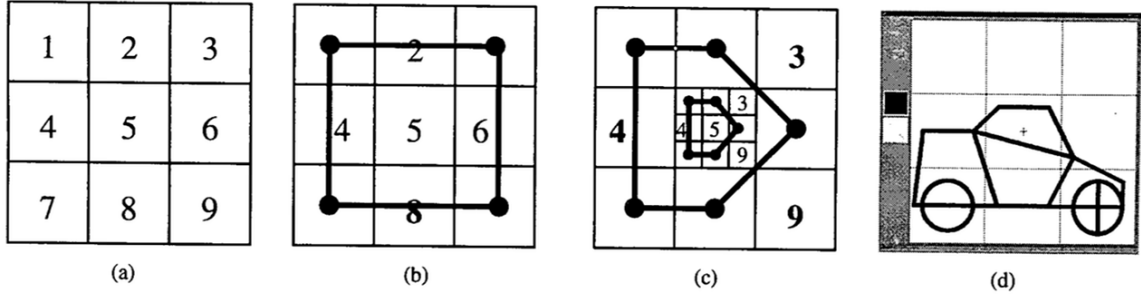


Figure 1. IC2D Model and Drawings: a) a 3x3 grid of cells, b) a square made up of lines, c) an arrow drawn inside an arrow using grid recursion, d) a car drawn by a participant in the IC2D pilot study.

Figure 9: Grids in the Integrated Communication 2 Draw. From [?]]

study of tactile drawing, , Kamel and Landay [?]] state that a drawing tool should allow following to be discovered: a point in relation to the drawing, the length of a line, curvature, and angles¹³. In all cases, IC2D provides access to these components with navigation in the grid and speech output.

For the purposes of graphs in K-12 education, a system like the IC2D has several drawbacks. One goal of IC2D was to “draw recognizable figures,” [?]], and certainly the car in figure 9 is recognizable visibly. However, it does not follow that the car would be recognizable to other blind people or even the blind author at a later time. The initial system gave verbal point feedback, but no overview of what was drawn. While labels provide more context [?]], there is a limit to descriptive power, they take time to develop, and are based on the describer’s expectations of what the observer wants to know. In addition, there is a cognitive element that appears different. If the purpose is a description, then a drawing application may not be the appropriate means of transmission. The picture itself may have meaning that is not easily converted into words, or there are so many words to describe the picture that a description becomes too long. While the IC2D excels at point estimation, it may not provide sufficient trend analysis and overview for graphing tasks.

The second challenge is the use of a grid. Using a recursive grid is a powerful tool, leading to logarithmic time complexity. Given a 1-dimensional grid of 3 points, a user can

¹³While the selection of these four properties appears tailored for the IC2D, certain ones do appear to have external validity. Finding a point and finding the length of a line are both part of the sixth grade graphing standards, specifically CCS 6.NS.6.c and 6.G.3, from [?]].

recurse into any pixel in a pixel space of w within this number of key presses:

$$k = \lceil \log_3(w) \rceil \quad (4)$$

In two dimensions, with a pixel space of 1000 x 1000 pixels and a 3x3 grid, any point can be targeted with a maximum of $\lceil \log_3(1000) \rceil = 7$ key presses! This is a phenomenal performance, but there are critical drawbacks. First, a pilot for the IC2D found that participants had trouble with more than 3 levels of depth [?], leading to a maximum of $3^3 = 27$ points for horizontal and vertical directions, or $27^2 = 729$ cells. While this may be suitable for recreational drawing, it is simply insufficient for graphing. A standard domain and range of -10 to 10 has 21 tick marks, so with 27 points per dimension, such a system could not even represent the halfway point between tick marks. For a 27x27 grid overlaid on a 1000x1000 pixel space, each direction of the cell would represent over 37 pixels, or about 0.35 inches on a typical screen. The IC2D would not provide enough accuracy for a graphing problem such as “Plot 2.4 on the number line,” (a sixth grade graphing problem based on the CCS). The grid has very different movement than the mouse as well, with unclear results in terms of conceptual understanding. In many cases, for example, two cells will be spatially adjacent but require a very different set of keys to reach.

Third, and most importantly, the speech output and interaction structure is much different than what sighted students would have. From a construct perspective, this is problematic: blind students are not learning the same principles as sighted students. In addition, IC2D may give more information than what a sighted student would have available. For example, if the student were asked to plot the point (2,3), the blind student would be listening for the labels 2 and 3 at a particular point, then select that point. In visual graphs, sighted students would look for tick marks, and their associated labels. Following similar interaction behaviors as sighted students could support use of the same cognitive constructs¹⁴.

Kamel and Landay’s suggestion that the keyboard can replace the mouse is interesting

¹⁴Kamel and Landay [?] report that many blindfolded and blind participants reported that they could understand the spatial characteristics of the figure, and have success creating accurate and detailed drawings. Perhaps the recursive grid can be used to maintain the construct, but since the behavior is so different, it appears that the answer lies in further research.

on a number of levels. First, the keyboard is more heavily used than the mouse by visually impaired populations. Fitts's Law studies show a logarithmic targeting speed with the mouse, like the keyboard movement used in IC2D¹⁵. However, the feedback mechanism for the keyboard is speech, while it is spatial for the mouse¹⁶. Due to the relatively long processing time for speech, the user attempting the keyboard is likely to take much longer during each step, essentially reorienting each time.

In conclusion, the IC2D provides an efficient, accessible means to plot points in a 2-dimensional space. With a keyboard and speech output, blind people can find and create objects. However, in terms of grid size, keyboard interaction, and alignment with K-12 graphing, the IC2D would require further adaptations.

2.3.3 Upson's Education Evaluation

Robert Upson was the first to move sonification squarely into the educational space. In the early 2000's, he explored the use of sonification for students learning mathematics [?], and evaluated his Sound Grid [?]. Upson emphasized the use of multiple representations for mathematical concepts:

[...The cited] studies show that presenting mathematical concepts in several representations at once can improve understanding. Interrelationships may develop among representations presented concurrently. The student may also be taking ownership of his learning by having a choice of representation.

Upson thought of sonification as a supplement to visual graphing, as a way to learn more about the graph interpretation. Participants in his study were elementary and middle school students, in standard and alternative¹⁷ schools.

¹⁵The mouse is generally faster than the keyboard for sighted users [?] and slower for blind users [?], but the IC2D adaptation may give the keyboard an advantage for both users. Phase 1 of this dissertation (Chapter 4), shows how the mouse can be much faster than the keyboard given certain non-speech feedback.

¹⁶Of course, the user could see the screen and get spatial feedback with the keyboard. This was explored in Phase 1, study 2, with sighted users (Section 4.2), with a different type of keyboard input than found in IC2D. Sighted participants were much faster with the mouse in the visual, auditory, and combined conditions, similar to visually impaired participants with auditory feedback in study 1 (Section 4.1).

¹⁷Upson was primarily concerned with alternative representations that help people learn math, which motivated his use of alternative schools (from [?]):

Alternative schools serve students who have not fit traditional classroom settings. Losses of

Upson’s research began with teachers. In two presentations, with a high school science group and a high school math group, Upsen played sonifications of weather data. In the first presentation, from the “twenty-five participants [...], only one inquired further [...due to her] visually impaired chemistry student,” [?]. The apathy turned into negative comments at the second presentation, where the sonification was deemed “disruptive to class”.

The positive outcome of the teachers’ feedback was Upsen’s shift of attention to learning standards. In interviews [?] and classroom activities [?], Upsen presented graphing questions that were loosely based on standards, or taken directly from previous Oregon State tests. Upsen added opportunities for students to experiment creating graphs themselves, with a piano, pencil and paper, Microsoft Excel, and a custom interactive software application called SoundGrid [? ?]. Unlike many other auditory graph evaluations in education, student participants were fully sighted and could use both their eyes and ears.

Upson’s Sound Grid presents a two-dimensional space where students can plot points (Figure 10). The scale of the tick marks and auditory graph settings are managed through a control panel next to the graph. Students can hear the graph, with the y values mapped to distance and the x values mapped to time¹⁸. Settings such as note length and instrument could be used to perceptually group points. Sound Grid also featured a data table for data entry and viewing similar to a spreadsheet [?].

In both studies, results were mixed. Student feedback was generally positive, and used musical analogies to explain the graphs. Students appeared to understand the mapping concept, but basic concepts such as x and y axis and the specific mapping were often lost [?]. In the more formal test, Upsen found an increase in answered questions in the post-test, but most of these answers were incorrect. For the non-equation questions, about 10% were

family members, pregnancy, learning disabilities, and behavior problems (including criminal activity) are some reasons for enrollment at alternatives schools. To some degree, the participants in this study have demonstrated the failure of traditional classrooms to serve their educational needs.

Note that “alternative” schools are for alternative approaches for education for a variety of reasons. They are not “specialized” schools, such as a school for the blind, which focuses on a particular set of approaches. Specialized schools are an important resource for understanding are discussed in more detail in Section

¹⁸While not stated specifically, it appears that the x values were mapped to time, since playback rate and note duration were variables that could be managed. In addition, X values could optionally be mapped to the volume, an unusual choice that is less likely to be as useful as pitch [?].

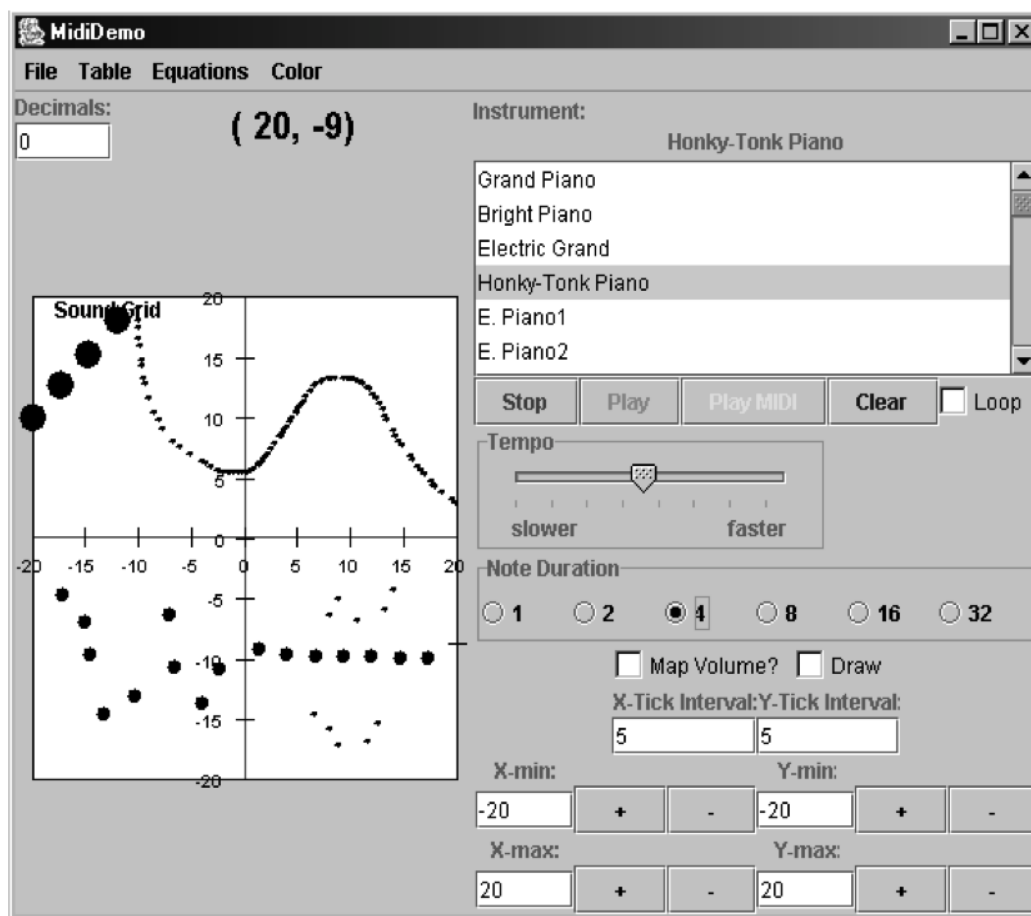


Figure 10: Upson’s Sound Grid, featuring an interactive, two-dimensional canvas, and settings for auditory graph playback. From [?].

correct, 50% were incorrect, and 40% were not answered. For the equation questions, about 5% were correct, 20% were incorrect, and 75% were not answered¹⁹. In

These results are astonishingly low. Upon reflection of Upson’s strategy, a few challenges arose that may be avoidable in future studies, specifically participant selection, the relationship between music and sonification, and specificity of graphing problems. Upson was “using sonifications as supplements to graphing,” [?]. Upson reasoned that alternate graph formats would help students learn the content, so he also wanted to help students who have learning challenges, such as those in alternative schools [?]. Sighted students, however, may not gain much from sonifications. The auditory display was impoverished:

¹⁹I sympathize that these quantities are wanting. The papers display the unfortunate aspect of graphs, that of estimation. Many quantifiable specifics were not reported, as Upson focused his report [?] on qualitative student performance and feedback.

recreating the graph in figure 10 only given a sound playback would be impossible. Reading a visual graph is quick, and sighted students may become frustrated with the slow process of listening to a graph²⁰. The students in Upson's second (classroom) study had initial graphing scores that indicated they did not understand the material. Therefore, Upson had to show that the lesson with Sound Grid provided both an alternative format and a tool for learning graphs within a small time period. Instead of using sighted participants, alternative or mainstream, perhaps auditory graphs would be more suitable for blind students, who have *no* computerized access to graphs. This is supported by the sole interest from the teacher who had a visually impaired student. Since such a system would require completely non-visual accessibility, perhaps some of the auditory shortcomings could also be resolved.

The lesson featured sonification, but confounded music with data analysis. Many sonification terms are also used in music, but the goal of sonification is data analysis, which sometimes has different outcomes than sounds that are more aesthetic or musical. The terminology Upson uses shows an excessive overlap (from [?], emphasis mine):

Honestly, algebra was of little use to the students for *composing*. The concept of *melodic* contour in linear ascending and descending series of *notes* created by algebraic equations was too complex and sterile. The Sound Grid allowed participants to freely enter *notes*, sometimes in lines, sometimes not.

Upson appears to want students *composing* graphs²¹. But the purpose of graphing is data analysis, often a sterile practice when compared with music-making. While musical concepts can be adapted for sonification, in educational contexts, there are no aesthetic requirements for a student's graphs beyond simple legibility. It is much more important to be able to read and write simple graphs.

²⁰The differences found in Phase 1 Study 2, Section 4.2, show visual graphs to be about 3 times faster. In the pilot for Phase 3 Study 1, Section 6.1.2, two low vision students who were competing for speed on graphing questions had many acceptable answers when they used vision, but skipped or erred when they then used auditory graphs. These students could use the auditory graphs, but, I assume, they gave up when they felt the auditory format was slowing them down.

²¹I believe Upson understood that sonifications are about data analysis. However, the way sonifications were presented to students and discussed in the paper was often as music. Upson probably decided on this approach to make graphs more interesting, but this made the task more difficult, since the goals of music and data analysis are so different. The scientific and artistic aspects of sonification are represented in publication venues such as the International Conference on Auditory Display (ICAD).

Upson was the first to work with teachers. This led to an attempt to tie the lesson plan with the learning standards, and one-to-one and classroom interaction with K-12 students. All three parts are critically important for practical graphs, but could be improved. First, as mentioned above, the biggest gains may be for visually impaired students, particularly those who are severely blind. Their teachers, in turn, may be more inclined to use sonification (as suggested by the teacher feedback in [?]).

The use of standards is more complicated. Upsen is an educator, and knows the art of lesson design in line with state standards. However, for many software developers, it may be unclear how to tie standards to specific lessons. A more systematic approach may be necessary. This approach, if extended well beyond Upsen's intention, may also be used to create the design requirements for the sonification system. This dissertation defines such an approach, called SQUARE, in Phase 2, (Chapter 5)

Upson presented a graphing system that provided sighted students a canvas for plotting points. He moved sonification into the classroom. Yet much more development was necessary before graphs could be used for practical purposes.

2.3.4 Audio-Haptic Graphs in MULTIVIS and related projects

Over the past 12 years, Stephen Brewster's group and collaborators have conducted a large body of research related to accessible graphs, spanning user requirements, psychophysical studies, system development, and evaluation. The analysis of this work is designed to summarize key findings of the program of research, and is divided into three sections: users, system, and evaluation.

2.3.4.1 Users

This research started in the classroom. Ramloll et al. observed teachers and their visually impaired students in a mock classroom [?]. The tasks were for students to a) label axes on swell paper; b) find points on raised grid paper; and c) create graphs with pins and rubber bands. Ramloll et al. commented on the predominance of tactile tools. They were particularly concerned with student dependency on teachers for the tasks:

We are also motivated to achieve user independence by designing a system that

will not require the intervention of a sighted helper in order to be used. It is also important that our system encourages users to focus on data comprehension tasks rather than on the construction of the data representations, as is the case with offline printing using embossers and rubber band graphs [?].

Later papers motivated the use of multimodal graphs in terms of challenges with current tactile technologies. Some suggested further work [?] or conducted similar analyses [? ?]. Many evaluations involved visually impaired participants, some of whom were students. However, it appears that the program of research has yet to close the loop, and evaluate in classroom situations.

2.3.4.2 System

The graphing systems focused on independent graph creation and exploration, likely due to the motivation from the user requirement studies. Ramloll et al. emphasized the use of active graph representations, suggesting they are more suitable for understanding and using graphs:

[I]t is necessary to distinguish between the passive representation of line graphs and their active representation [...] in the former the reader is a passive receptor of information while in the latter the reader is in control of the flow of information. We postulate that putting the reader in control of information access may have significant effects on the sustainability of reader-interest and arguably on the retention of the information accessed. [?]

The authors show an interest in providing the user with controls that can be used to filter the information presented to specific parts of the graph.

The systems presented varied widely, but often had audio and haptic components. A early novel system presented a y-pitch mapping and an x mapping to time, volume, and pan (using a sound source model) (Figure 11) [?] (and extended in [?]). The display of audio often mapped the y-axis to pitch (based on their own guidelines [?], but the particular mappings of both x and y depended on the publication, along with the type of

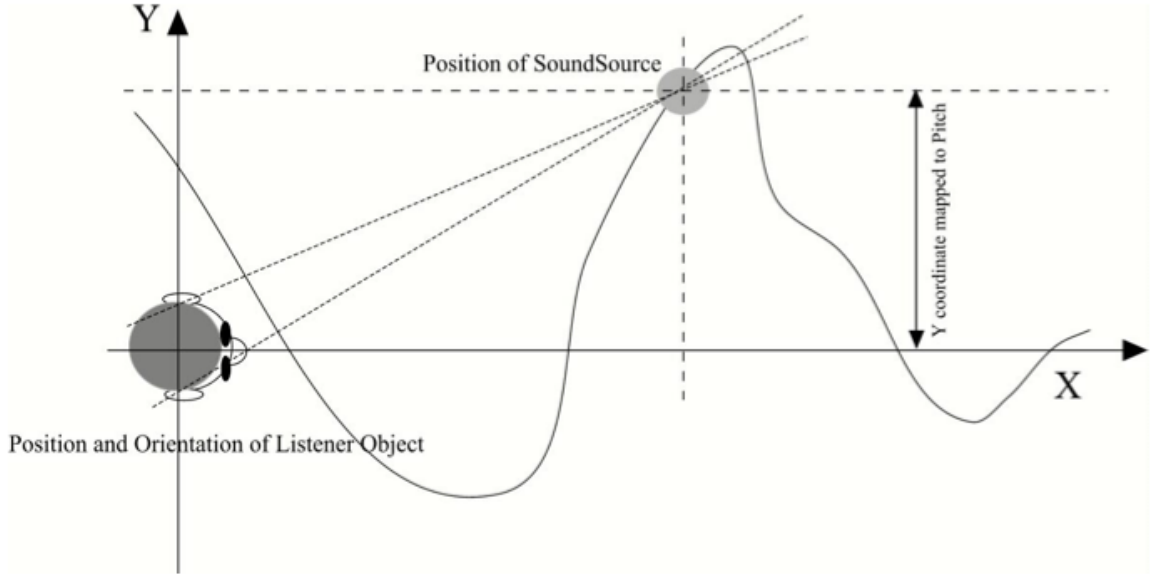


Figure 11: Auditory display used in Ramloll et al. [?]. Pitch is mapped to y values. Volume and pan are mapped in a way to simulate what would be heard from a sound source if the listener were standing at the origin, facing the positive direction of the x axis.

Table 4: Key features of Figure 12, as defined by [?]. Note the emphasis on trends.

| Line 1 (black) | Line 2 (gray) | Interaction |
|---------------------------|----------------------------|----------------------------------|
| 1. Curve | 5. Straight line | 7. One intersection point |
| 2. Generally decreasing | 6. Continuously increasing | 8. Intersection at halfway point |
| 3. Level section at start | | 9. Maxima roughly equal |
| 4. Slight increase at end | | 10. Minima roughly equal |

interactivity. MULTIVIS emphasized the use of multiple views, which included subsystems such as GraphBuilder [?] and SoundBar [?].

Early work in haptic research explored the use of computer vision to convert visual paper graphs to haptic graphs [?]. Challenges to point estimation and line identification for haptic graphs were identified early on. For example, for haptic grid lines distracted users from data lines, overlap data lines, only provide approximate values, and counting grid lines (grid lines were unlabeled) [? ?]. Speech and non-speech audio were proposed as a means to bypass the problem. In many cases, the systems depended on both haptic and audio to provide the entire scope of information.

Data input devices varied, but often involved a PHANTOM haptics device. Keyboards were sometimes used for navigation (single step only) [? ?], and often for data entry. The

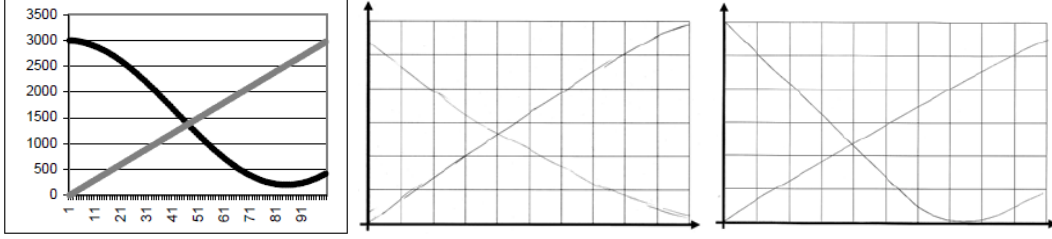


Figure 12: An example of graphs from a 2003 evaluation, “Drawing by Ear” [?]. The left-most graph is the original data. The middle and rightmost graphs are participant drawings of the graph, after listening to an auditory graph of the data. Note the differences in the trends and lack of verbal context. In one aspect, trend analysis, the drawings look similar to the original. However, without points, it is impossible to know the scale of the resulting graph. Participants were not given verbal cues for the task, which probably hindered their construction of the graphs in the first place, and led to minor errors in the trends.

mouse was determined to be inappropriate, due to feedback about lack of experience with the mouse [?]. Plotting points wasn’t common, but when it was, it involved a data table in early iterations [?], and absolute positioning or PHANTOM devices throughout the research (e.g. [? ?]). The major obstacle reported for the PHANTOM haptic pen is the price, which during the early studies cost about \$10,000[?]. However, it was shown that participants using cheaper force feedback mice (about \$60), when combined with audio, could achieve similar accuracies as with the PHANTOM. In addition, the PHANTOM only provides one point of interaction, like a finger, whereas more time efficient and realistic interaction involves several fingers [? ?].

2.3.4.3 Evaluation

Evaluations related to graphs can be grouped into lines, points, and combination. For lines, participants were asked to identify a line, describe the trend, indicate whether a trend is increasing or decreasing, count the number of lines, and count the number of bends in the line. The line evaluations required trend analysis, but no context or point estimation. For points, participants were asked to find the largest or smallest value²², find the two closest y values, compare y values, and find a specific value²³. In most cases, point selection was not

²²“Locating” the value involved finding it with a mouse or keyboard interaction.

²³This value was requested by the user, and spoken. It was a component of the data itself and not the context. That is, the data’s value was completely reported, as if the data were looked up from a table, not like an inference from tick mark values.

Table 5: Graph component aspects of evaluation, throughout the 2000-2010 research in multimodal graphs in the Brewster lab. Table is broken by line, point, and hybrid evaluations.

| Topic | References, in chronological order (2000-2010) | | | | | | | | | | | | | | | | | | | |
|---|--|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|------|------|
| | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] | [9] | [10] | [11] | [12] | [13] | [14] | [15] | [16] | [17] | [18] | [19] | [20] |
| Find the line | | | | x | | | | | | x | | x | | | | | | | | |
| Find the number of lines | | | | x | | | | | | x | | x | | | | | | | | x |
| Find the number of bends in a line | | | | x | | x | | | | x | | x | | | | | | | | |
| Find the trend | | | | | | | | | x | | | x | | | | x | x | | | |
| Find increasing or decreasing trend | | | | | | | | | | x | | | | | | | | | | |
| Find the largest or smallest value | | | | | x | | | x | x | x | | x | x | | x | x | x | | | x |
| Find the two closest y values | | | | | | | | | x | | | x | | | | | | | | |
| Compare y values | | | | | | | | | | | | | | | | | x | | | |
| Find the specific value | | | | | | | | | | | | | | | | x | | | | |
| Find the point at which two lines intersect | | | | | x | | | | | | | | | | | | | | | x |
| Find the visual graph from a list of options, after being presented an auditory graph | | | | | | x | | | | | | | | | | | | | | |
| Performance compared with tactile graphics | | | | | | | | | | | | x | | | | | | | | |

point estimation, since there were insufficient cues (specifically verbal cues) to give a value for the point. Instead, point selection involved using an input device to move to the point. For combinations, participants were asked to find the point at which two lines intersect, identify the visual graph after being played an auditory graph, and using both multimodal graphs and tactile graphs. Table 5 provides an overview of the use of the different graphing evaluations from publications authored by Brewster and other coauthors.

Most of the evaluations focused on reading graphs. However, participants were evaluated on graph creation in the Graph Builder and Tangible Graph Builder publications [? ?]. The first study involved typing in the range and domain of the graph, along with plotting the specific x (categorical) and y (ratio) values for a bar chart. Feedback after point entry, however, did not provide specific values (or context), and the study found many off-by-one errors [?]. The second publication evaluated the construction of bar and line graphs [?].

The body of research had a variety of other evaluations. Almost all measured proportion correct²⁵ and specific changes to the interface. Other evaluations included task completion time, user interface questionnaires, perceived difficulty²⁶, NASA TLX for cognitive load,

²⁵What constituted correctness depended on the study. Table 5 provides a list of examples.

²⁶The “perceived difficulty” was not compared with other formats, such as visual or tactile graphs.

Table 6: Other aspects of evaluation, throughout the 2000-2010 research in multimodal graphs in the Brewster lab.

| Topic | References, in chronological order (2000-2010) | | | | | | | | | | | | | | | | | | | |
|------------------------------------|--|-----|-----|-----|-----|-----|-----|-----|-----|------|------|------|------|------|------|------|------|------|------|------|
| | [1] | [2] | [3] | [4] | [5] | [6] | [7] | [8] | [9] | [10] | [11] | [12] | [13] | [14] | [15] | [16] | [17] | [18] | [19] | [20] |
| Task completion time | | | | x | | | | x | | x | | x | x | | | | | | | |
| User interface questionnaire | | | | x | | | | | | | x | x | | | | | | | | |
| Perceived difficulty ²⁴ | | | x | | | | | | | | | | | | | | | x | | |
| NASA TLX (cognitive load) | | | | x | | x | | x | | x | x | x | | | | | | | | |
| Graph sketches | | | | x | | | | | | | x | | | | | | | | | x |
| Log analysis of pointer positions | | | | x | | | | | | | x | x | | | | | | | | |
| Think-aloud graph exploration | | | | | | | | | | | | | | | | | x | | | |
| Conversation analysis | | | | | | x | | | | | | | | | | | | | x | |
| Interviews | | | | | | | | | | | | | | | | x | | x | | |
| Impact on collaboration | | | | | | x | | | | | | | | | | | | x | | |

graph sketches, logs of pointer positions, think-aloud graph exploration, conversation analysis from recordings, interviews, and the impact on collaboration.

2.3.4.4 Discussion

The range of studies in the body of work from Brewster and his coauthors cover an impressive, vertical exploration of graphs, from psychophysical studies to system evaluation. Given twelve types of graph task-oriented evaluations shown in Table 5, it may appear that graph literacy is largely covered. However, in the initial paper [?], the three tasks the teacher asked the students to conduct were much more centered on point estimation topics: label axes, find points, and create graphs. It may be more suitable to focus on point estimation tasks. However, in some cases, the research group goes to lengths to avoid text and context:

The participant was presented with a tactile raised paper bar chart, consisting of axes and seven bars. No labels, titles or legends were provided so as not to distract the participants from focusing on the tactile representation of the bars [?].

Some of this may be controlling for experiments, controlling the stimulus for the purposes of the study. However, point estimation may be more basic than trend analysis, so perhaps the experiments are starting at a more complicated level of graph literacy.

More generally, the selection of the graphing goals appears to be largely arbitrary, with a couple exceptions. McGookin and Brewster [?] leverage a theory of graph comprehension

[?] to build their questions²⁷, but the relationship to actual classes is less clear. Their later Graph Builder publication [?] uses questions from Scottish education tests. In these more grounded evaluations, point estimation appears to play a larger role. However, it remains unclear to what degree of coverage the evaluations have over the broad activity of graph literacy.

Related to point estimation, is the use of speech. While the group found participants preferred speech, it tended to be slower than other approaches [?]. Many of the displays had no verbal components (text, speech, or braille), so participants could not easily find or estimate the values (this sometimes led to off-by-one errors [?]). Some displays did have verbal components, such as speech in [?]. A specific study on the impact of speech and pitch was reported in [?]. Oddly, the lab’s guidelines for including speech extend to tables [?], but are not addressed with respect to graphs.

The “active graph” argument put forward in Ramloll et al. [?] has interesting implications for auditory graphs²⁸. If the “playback” method of graphs, where the x-axis maps to time, then an observer of the interaction could conclude the user is receiving the graph passively. While they are clicking a “play” button, they cannot control the information flow (it is managed by the x axis data, and the time given for playback). In order to hear one piece of the graph, the entire graph must be presented. Beyond the cognitive implications, it may simply be faster for a user to find what they are looking for when the information target and flow can be directed by the user. In short, time may not be an appropriate mapping for x values²⁹. Time is then an artifact of interaction with the graph. The question is then, of course, what should be used to map the x axis values.

²⁷McGookin and Brewster are, unfortunately, not clear on what exactly was asked, giving only one example: “The following graph shows the number of people who said that a particular type of food was their favourite. What were the three most popular foods?” [?].

²⁸While I argue traditional x-time mappings lead to “passive” graphs, referring to Ramloll et al. [?], it is important to note they used time as an x-axis mapping. They could have followed their argument further with respect to the auditory graphs presentation, but they may not have since haptic technologies were additionally available for the graph interaction.

²⁹This is not to completely object to the time mapping for x or other dimensions. In certain cases, such as an overview, a quick playback may be the most suitable. However, the selection of mappings should not be canon; it is based on practical needs.

The selection of this mapping can be motivated by number lines, which are one dimensional graphs. While not explored in previous accessible graphs research, number lines are prevalent in early K-12 education. A presentation of number lines as an auditory graph could play all the points at once, as a “chord” of sound. This may be difficult to understand, particularly with additional tick marks. It could also be a constant pitch, which plays notes on an x-time mapping when there are points associated with the current time value. In this second display, for interactivity, the user could move left and right with a device to indicate a position on the number line, and a tone could play back indicating the presence or absence of a value. Pitch is fixed, and could be used to map something else, such as distance to the nearest point. This would lead to an understanding of where a point is on a number line, with respect to the current position and fixed anchors such as tick marks. In other words, a pitch-to-distance mapping on a number line would provide point estimation, one of the most difficult elements of auditory graphs. The speed and accuracy of distance mapping for number lines in terms of target width, target distance, and audio mapping type was evaluated in Studies 1-4 in Phase 1 (Chapter 4).

Both the pitch and time elements, then, have been un-mapped from x and y values. Time has become an element of the interaction. Pitch indicates the distance to a target. In two dimensions, the one dimension approach can be duplicated. In other words, the distance-pitch mapping can be created in a *second* sound. The combination of pitches, after Phase 1 evaluation showed its efficacy, was used as a key component to the Graph and Number line Interaction and Exploration system (GNIE) system.

2.3.5 General Discussion

All of the previous examples of accessible graphs were motivated by improving K-12 education. However, there appears to be insufficient appreciation of current K-12 practice. An important step before graph evaluation would be to identify the specific evaluation goals, with respect to education goals.

A second point is to provide access, but limit the power of the system to the level which other students will be using graphs. This may be with interaction and navigation, such as

selecting a relational movement system instead of the hierarchical drill-down. In addition, a new tool should not reduce the cognitive task beyond that of what a sighted student would experience. This consideration is based on testing accommodations, which will be covered as part of the background on current practice in the next section.

2.4 Current Practice: Curriculum, Graph Literacy, and Visually Impaired Students

This section describes standards, curriculum, graph literacy, and the application of all three with visually impaired students. “Curriculum” is a term covering many aspects of education delivery. United States education departments have increasingly grounded curriculum in standards, currently using state requirements such as the Georgia Performance Standards (GPS), and moving toward the interstate Common Core Standards (CCS). A standard guides curriculum, and within a curriculum are specific tasks, such as plotting a point on a graph. This goals-components model has been championed in education for decades, guided by Tyler’s seminal “Basic Principles of Curriculum Instruction,” [?]. It is also present in Human-Computer Interaction (HCI), with models such as Goals, Operators, Methods, Selectors (GOMS) [? ?] and task analysis [? ?], which all have a clear hierarchy of tasks, broken into manageable operations that can often be evaluated against alternatives. Yet graph literacy, particularly for visually impaired students, is not well understood on the level of operations. The first step to make electronic assistive technologies for visually impaired students is to identify the operations that the student will have to complete. Furthermore, in the context of testing accommodations and fair alternatives, visually impaired students should be given a tool that gives them the opportunity to graph, but does not give them an advantage over their fully sighted peers.

2.4.1 Understanding Curriculum

When exploring the education space, the key material that is experienced by the student is the curriculum. A “curriculum” can mean many things. Bobbitt’s original definition was twofold: the formal and informal experiences a student has in and around school, and the “series of consciously directed training experiences that the schools use for completing and

perfecting the [learning process],” [?] (also see [?]). Our concern for this dissertation is formal learning, encompassing the “learning experiences” [?] students have while using the the tools, textbooks, test questions, homework assignments, learning environments, and method of instruction. With still such a broad definition, how can a curriculum explain the current learning situation?

Bobbitt’s approach was practical [?]. First, identify major occupations and the skills required. Then, discover what students have not learned of those skills through informal training. For example, many children learn language at home, but still make grammatical errors such as double negatives [?]. Based on known deficiencies, design a series of lessons to overcome those shortcomings. Students will then have learned the necessary skills to overcome the “social shortcomings” in the area of study. Bobbit presented three important themes that were later adapted by Tyler [?] and are present in the United States education system today [?]: curriculum-making as a scientific process, curriculum defined in terms of formal education experiences, and a regional view of learning outcomes (larger than the classroom or school).

As a tool for understanding shortcomings of the Bobbit-Tyler scientific perspective, it is useful to listen to the many critics of the approach at the time [? ? ? ? ?]. Montessori cautioned on the overuse of cold and calculating methods, showing how the results were not always in the students’ best interests [?]. Montessori also championed the teacher as an architect of customized lessons, based on student needs. Dewey echoed similar concerns [?]. Both influential educators supported science, but saw the modularization of education as a dehumanizing force. We, as parents, educators, and citizens, might all have an internal conflicting view of student and teacher independence alongside general standards and curriculum materials. These concerns could be addressed within the course of a comprehensive evaluation. Such a study of the effectiveness of a graphing tool should capture both the examination-oriented alignment with standards, alongside a field study of the tool in rich [?] classroom settings.

Bobbitt’s work was extended in the 1960’s by Ralph Tyler. Tyler’s solution was to identify the intended goals, and build a learning system that best enables students to

meet those goals. In his seminal work, “Basic Principles of Curriculum Education,” Tyler provided four questions to ask while designing a curriculum [?, page 1]:

1. What educational purposes should the school seek to attain?
2. What educational experiences can be provided that are likely to attain these purposes?
3. How can these educational experiences be effectively organized?
4. How can we determine whether these purposes are being attained?

This dissertation, however, is not one on curriculum studies. My focus will be on item 2, the educational experiences. Item 1, the educational purposes, are high-level goals that students are expected to learn in some formal learning environment. In Georgia, high-level education goals are defined by the Georgia Performance Standards (GPS), and in 2014, the Common Core Georgia Performance Standards (CCGPS). Item 3, the organization of educational experiences, is the role of the curriculum developer and the teacher. Item 4, evaluating student experiences, involves multiple forms of assessment.

Tyler defines “learning experience” as [?, page 63] (original emphasis):

the interaction between the learner and the external conditions in the environment to which he can react. Learning takes place through the active behavior of the student; it is what *he* does that he learns, not what the teacher does.

Tyler includes a list of principles for the selection of a learning experience. As this dissertation suggests a method for creating education technology, an understanding of components of the learning experience will assist in the development of evaluation instruments. The following five quotations explain Tyler’s view on the selection of learning experiences, and potential application to an accessible and educational graphing technology.

2.4.1.1 Selecting learning experiences

“[A] student must have experiences that give him an opportunity to practice the kind of behavior implied by the objective,” [?, page 65]

Tyler emphasizes practice, with materials directly related to the topic. This is an important consideration for assistive technologies. As one goal is to build an alternative technology for graphing, then the tool must enable students to build and interpret graphs, in a way similar to how sighted students build and interpret graphs. Beyond *enabling*, students must actually *do graphing*, as defined by the objectives and the student-centered “learning experience.” More practice also leads to faster and more accurate execution of the learning experience. Thus, an assistive educational technology should, ideally, allow the blind student to practice all components of graphing, and complete about the same number of learning experiences in a set period of time, as his or her sighted peers. Time and accuracy, then, are critical components to know for both blind students and sighted students.

There is a comprehensive aspect of meeting objectives that must not be overlooked. First, the technology will ideally suit a number of objectives and learning experiences. For example, it may be suitable for most graphing objectives and graphing questions found in a Georgia middle school. A claim of being “suitable for middle school” is easily defended if the technology is systematically evaluated with each learning objective, through appropriate learning experiences. Practice must be available for each of these experiences. Therefore, a comprehensive evaluation will consider every learning objective (standard) the technology claims to support.

In addition, the precise approach to meeting the objectives must also be similar to traditional methods and the intention of the objective’s authors. In education, the technology must not overstep the line of providing support to answering the question for the student. Consider a graphing calculator. It enables students to quickly compare graphs, without taking the time to plot the points by hand. This useful technology, however, oversteps its role in the classroom if the task is to plot points. This “graphing calculator problem” is similar to the “calculator problem” with arithmetic. Based on this “graphing calculator problem”, a consideration of the steps the student is expected to know, and what the technology should *not* do for the student, are important aspects of the technology’s design.

“[T]he learnings experiences must be such that the student obtains satisfactions

from carrying on the kind of behavior implied by the objectives. [...] If the experiences are unsatisfying or distasteful, the desired learning is not likely to take place.” [? , page 66]

Students must, on some level, appreciate the material being taught. This is heavily influenced by the curriculum and the teacher. However, an assistive technology should not reduce the level of interest in a topic. If it increases interest, it may improve the opportunity for learning. For example, if students gain more interest in a lesson since they are using a laptop instead of paper, then the students may be more motivated to complete the learning experiences (practice), and eventually satisfy as much, if not more, of the objectives. Therefore, a comprehensive evaluation of technology will include measures of student interest through subjective questions, ideally comparing the new technology with current methods.

In practice, I have found two challenges with auditory graphing tools: speed and sounds. Graphing using vision usually takes less time than with sound. So low vision and sighted students who are forced to use a slower interface sometimes get frustrated. In addition, a handful of students who can use vision don't like the sounds, and want to take off the headphones. The solution I have found in these cases is to simply allow the students to solve the graphing questions in the way they prefer best. In addition, I document their use of tools. A comprehensive evaluation should specifically include subjective evaluations of the sound design and perceived speed.

“[T]he reactions desired in the experiences are within the range of possibility for the students involved,” [? , page 67]

This claim echoes the Zone of Proximal Development, where students can learn within a certain range of possibilities, based on what they already know[?]. For assistive technology, there is a second meaning, extending Tyler's intended definition. Visually impaired students have barriers on their sensory development of graph literacy, based on their lack of vision. This barrier will not disappear directly (they will not become sighted). Therefore, the “range of possibility” may never extend to visual graph literacy. However, visually impaired

students may learn a graphing standard by using an alternative format. Tactile graphics, for example, are an acceptable alternative to visual graphs in many cases; although tactile graphics do not require the development of visual skills, they do require the development of similar perceptual and cognitive components of graph literacy, such as relationships between two points, finding the origin, and discovering the graph family. An assistive educational technology evaluation, then, should show that students can conduct the same sort of tasks. This could be done by comparing the sort of errors that students make between formats, converting between formats, and asking students to describe the graphs they see, or hear, or touch.

“[T]here are many particular experiences that can be used to attain the same educational objectives,” [? , page 67].

The concept of several ways to learn the same concept is leveraged in education, and is a key concept in accessibility as well. Students learn graphs by reading graphs, writing graphs, watching others use graphs, and using graphs in basic ways and in applied ways. Using graphs in different ways allows a person to explore many facets of the tool, and the concept. Similarly, an assistive educational technology allows a student to learn the concept with a much different type of tool. It is critical, however, that the assistive technology can be used to learn the educational objective. This alternate graphing tool, for example, must not be a “graphing calculator” when “graph paper” is what is required.

Since the application of the tool for particular learning experiences is what has the student learn the concept, then an evaluation of the system could compare the students’ understanding and learning of the concept through different tools. While the experiences will not be identical, the student should be improving their understanding with practice on any tool. In addition, errors in understanding should be evident in both. For example, if a student does not know where quadrant II is on the graph, this should be obvious in similar errors with the original tool and the new tool. Therefore, system evaluation can have components of learning and accuracy.

“[T]he same learning experience will usually bring about several outcomes.

Thus, for example, while the student is solving problems about health, he is also acquiring certain information in the health field. He is also likely to be developing certain attitude toward the importance of public health procedures. He may be developing an interest or a dislike for work in the field of health,” [? , pages 67-68]

Tyler’s last point on learning experiences emphasizes the several outcomes possible, alone or in combination. The teacher may have a handful of these objectives in mind. one goal of STEM education is to increase the interest in the topic, and STEM fields. So, an evaluation of new tools for STEM education could also measure interest in the topic and future careers.

Tyler’s suggestions on learning experiences shape the development of assistive technology. Based on the five points, there a few key aspects for evaluation. The concept must be contained within the new technology. Evaluations should measure whether students understand and learn the concept, and whether the concept is being learned in a manner similar to standard technologies. A list of steps, like a task analysis, would make it easier to compare different tools. And finally, using mixed methods will capture both objective measures and the subjective sense of whether the tool is working.

The “experiences” and “organization” points suggest a nested structure of formal curriculum items. An experience could be on many levels, such as a daily lesson with several graphing problems, or a single subtask of a particular problem. Two of Tyler’s points were to make curriculum development a scientific endeavor, and to make the results of changes measurable[?]. Component parts could be measured, and compared with other approaches[?].

Analysis of nested structures is not new in Human-Computer Interaction, and I may be able to leverage HCI techniques within the education space. Hierarchical task analysis use the “systematic decomposition of goals and subgoals [...] to any desired level of detail,” [?]. GOMS and related methods like KLM [? ?] take this activity a step further, predicting speeds of certain human actions. For structured, serial actions such as plotting points on a graph, these approaches are well-suited. However, they are often a bit more detailed than

may be necessary for this particular task. A more lightweight approach, such as the Task Analysis for Problem Solving (TAPS) [?] may be more appropriate.

2.4.2 Common Core Standards

The educational purposes of a particular class are increasingly being managed at a level much higher than the classroom. In the United States, most schools are required to follow “standards” adopted by each state. Education standards are general requirements that a student is expected to know for a particular class. Georgia has the Georgia Performance Standards (GPS), for which there are standards for every grade of mathematics [? ? ? ? ? ? ? ?]. For the past several decades, classroom teachers have developed lesson plans based on state standards and district curriculum. Standards are intentionally broad and are independent of pedagogy. In other words, they can be learned by using several different ways of teaching, and several different types of tools. For example, a student can meet a standard by using visual, auditory, or tactile graphics, as long as the tools let the student learn the concept being taught. There are important caveats to this flexibility, covered in sections 2.4.1.1 and 2.4.3.

In 2014, all but two states will adopt the Common Core Standards (CCS) and integrate it into their state standards (technically, the GPS will become the Common Core Georgia Performance Standards (CCGPS)). The CCS are a shared set of standards between participating states. The CCS are currently being pilot tested in a few states, including Georgia. In the 2014-2015 school year, participating states are expected to have curriculum guided by the CCS. The CCS may be extended by the states to include up to 15% more standards, such as the CCGPS. For the grade targeted in this dissertation, sixth grade, there is a high overlap between the GPS and the CCS, so many curriculum components that are being used to meet the sixth grade GPS may be suitable for the CCS.

There are several reasons to adopt shared standards between states. First, students will have a shared set of expected knowledge that can be transferred to other states. This could assist in transition to college. Second, teachers will have similar concepts to teach between states, making their skills more transferable. Third, acceptable accommodations on tests

may be easier to determine with shared objectives, making it easier for administrators to determine acceptable accommodations for disabled students.

To understand standards, a more detailed example may help. A standard is a high-level criterion of student competence. For example, Common Core Standard 6.NS.7.c is “Understand the absolute value of a rational number as its distance from 0 on the number line; interpret absolute value as magnitude for a positive or negative quantity in a real-world situation,” [?]. This standard is a requirement for students to understand absolute value and its relationship to magnitude, and a requirement for students to represent absolute value on a number line. This standard with the other standards from grade 6 are the list of criteria that a student must know to be considered completely competent in grade 6 mathematics. While an educator may infer certain ways of teaching or assessing the standard, the standards themselves leave open to the instructor the particular way of learning the material, methods for solving the problems, or media to use. On a between-classrooms level, student competence of standards is typically measured in standardized tests in the United States. For this thesis, standards, particularly Common Core Standards (CCS), will be the “purposes” of the curriculum.

The learning experiences, described in section 2.4.1.1, include the specific graphing problems that they do, and the tools that they use. Specific graphing problems are often included in curriculum materials purchased by a school district. In 2011, my own research found over 700 graphing problems in the first half of a ninth grade mathematics textbook [?]!

Consider CCS 6.NS.7.c again. The first part, “understand the absolute value of a rational number as its distance from 0 on the number line,” appears very straightforward, but leaves open several aspects of teaching. For example, the media for the number line could be in paper or on a piece of software. A teacher could focus on integers, fractions, or decimals, and could have numbers on tick marks or in between. Assessment could include multiple choice questions, plotting on a number line, or verbal answers. Thus, standards are the required learning components, and curriculum is the way the standards are taught.

This dissertation will explore the use of graphs with the use of the CCS and to some

extent the GPS. Curriculum Changes will be focused on changing the medium of delivery from paper-based tactile graphics to computerized auditory graphs; otherwise, I assume that the curriculum will remain the same. In this respect, my dissertation is not fundamentally about changing education. It is about evaluating education technology (specifically auditory graphs); whether the technology can provide similar utility in otherwise the same curriculum and standards. By “utility”, I mean retaining the steps in the process as much as possible for students to solve the problem. Since curriculum accommodations could be used in testing situations, and since testing accommodations should have the students do similar steps, my approach may lead to software that could be used by visually impaired students in testing situations.

2.4.3 Testing Accommodations

Most students in the United States ultimately demonstrate proficiency in mathematics with an examination. While a comprehensive evaluation of the software in terms of testing accommodations is out of the scope of the proposed research, understanding the issue informs system design and evaluation. This section introduces testing accommodations, and the key issue of evaluating the same construct.

When altering the presentation or format of the test question, the construct being evaluated should stay the same. Alternative formats, such as calculators or text descriptions for graphs, can change the constructs being evaluated [? ?]. An alternative format, then, could use a metaphor of common formats as a template for design. With graphs and number lines, for example, tools could be based on the metaphor a pencil and graph paper or a graphing calculator; each approach would have tradeoffs.

If the constructs are the same for a task completed with one tool or the other, then an evaluation should demonstrate that those who have not had the opportunity to fully express their competency (impaired students) with the old tool should improve in their test scores with the new tool, while those who had sufficient tools (non-impaired students) will not improve in test scores; this is a common validity metric in testing accommodations [? ?]. I will apply these two heuristics to evaluate whether our system is an appropriate testing

accommodation. First, examination scores for visually impaired students should stay equal or improve with the new graphing system; a score reduction would indicate that proficiency with the new tool is harder to demonstrate than with the old tools. Second, examination scores for sighted students may stay equal or reduce with the new graphing system; a score increase would indicate that the new system is using a different construct than the problem intends.

2.4.4 Graphs and Number Lines for Visually Impaired Students

Graphs are a key component of K-12 mathematics education. Unfortunately, many visually impaired students cannot use standard graphs or number lines. This section describes visual impairment, visually impaired student learning environments, and tools used to read and write alternate forms of graphs.

2.4.4.1 Understanding Vision Impairment

It is tempting to define the term “visually impaired” based on medical diagnoses, for example using acuity and angle of vision. However, there are many other properties to consider, including type of impairment, other impairments, age of onset, and the effect on each eye. This list of properties could be collected, but in software development it is often more useful to consider the functional effects of the impairment. In other words, what is the *disability* stemming from the impairment? More specifically, can someone who is visually impaired use mathematics curriculum materials, including formulas, graphs, and geometry?

In this dissertation, “visually impaired” K-12 students are defined functionally into two categories, low vision and blind. “Low vision” students are those who use magnified visual graphs. “Blind” students are those who use tactile graphics.

According to the American Printing House for the Blind (APH), there were 59,341 students receiving federal funds in 2009³⁰ [?]. Of these, 5,411 use braille as their primary reading medium³¹, and are likely to also use tactile graphics. 16,075 use visuals as their

³⁰The American Printing House for the Blind (APH) is a valuable resource for information about visually impaired students in the United States. By federal mandate, the APH is required to register visually impaired K-12 students. States receive funds based on their number of students, so there is an incentive to count every student.

³¹More than the reported 5,411 many may be fluent in braille, but use auditory or visual methods, or are

primary reading sense, and are likely to use magnified visual graphs.

2.4.4.2 Mainstream and Specialized Schools

Most visually impaired students worldwide are in one of two learning environments: “specialized” schools for the blind, or “mainstream” public schools and classrooms with their sighted peers. In the United States, most visually impaired students remained in specialized schools until the Rehabilitation Act of 1973 [?], which encouraged the mainstream model. Since then, all states have increased their student participation in mainstream schools, but specialized schools continue to exist in most states. Currently, 83.2% attend mainstream schools, and 8.9% of visually impaired students attend specialized schools³² [?]. Many other countries also have specialized and mainstream schools, with a global trend toward mainstreaming.

Based on my own experience, there are several differences between specialized schools and mainstream schools. Students in specialized schools disproportionally have multiple disabilities and have more academic problems. Specialized school students are also more likely to be blind instead of low vision. Many students join specialized schools after mainstream schools fail to teach them basics, so specialized schools often have a much larger high school than elementary school. Specialized schools have most teachers familiar with both content and teaching to visually impaired students. Specialized schools also have centralized access to tools for the visually impaired, and act as a hub of information on this topic for the state. State meetings of vision teachers often occur at the specialized school, such as the Georgia Vision Educators Statewide Training (GVEST).

Unlike specialized schools, mainstream schools typically divide the roles of content teacher and vision teacher. Often, the vision teacher has several students to monitor throughout the district. Additional time is set aside during the day (about once per week) to train the student on technology and additional skills (such as the expanded core curriculum [?]). Mainstream schools must use more portable tools, and students may be more

not readers.

³²In addition, 5.2% are in work rehabilitation programs, and 2.7% are in multiple disabilities programs

sensitive to showing others that the student has an impairment. In summary, most visually impaired students are in mainstream schools. Expensive, noticeable, and non-portable equipment is particularly difficult for these students to use. It reduces the chances that the student will actually use the tool, which in turn reduces the chances the student will learn the material or the underlying standard.

2.4.4.3 Alternative tools

Tactile graphics are the standard way to display graphs to blind students. A tactile graphic is a raised impression of the figure to be displayed, and the accompanying braille information. Simple tactile graphics can be made with the 6 standard braille cells of a braille or slate and stylus; number lines are often created this way [?]. More sophisticated tactile graphics, or graphs saved on a computer, are printed out on an embossed sheet of paper, with more dots in an area than the 6-cell braille. The most complicated graphs are hand designed by experts and copied into a plastic thermofom mold. Tactile graphics are as old as braille, from the 18th century [?].

Creating tactile graphics is an art and a science. Often, the primary mapping is location-based, like visual graphs (see [?] and the discussion on graphicacy in section 2.1). However, tactile perception is different than visual perception, and the display cannot be converted from visual to tactile form in a simple manner. For example, it is difficult for people to follow a line with their fingers, when there are other lines within a few millimeters. Braille also takes up more space per letter than printed text. Thus, tactile graphics take up more space, and require a sophisticated modification of the original. When considering tactile displays for a computer, it is important to consider that there is no set algorithm for converting from visual to tactile graphics. Therefore, specialists create tactile graphics on specialized tools such as the Graphic Aid for Mathematics (Figure 13). They may also use embossers, which mark braille, raised points, and raised lines on paper. These embossers act as printers for commodity PCs. While this format is more convenient, before embossing the graphs are inaccessible to blind students, and the end products cannot be modified by students (they can only be read).

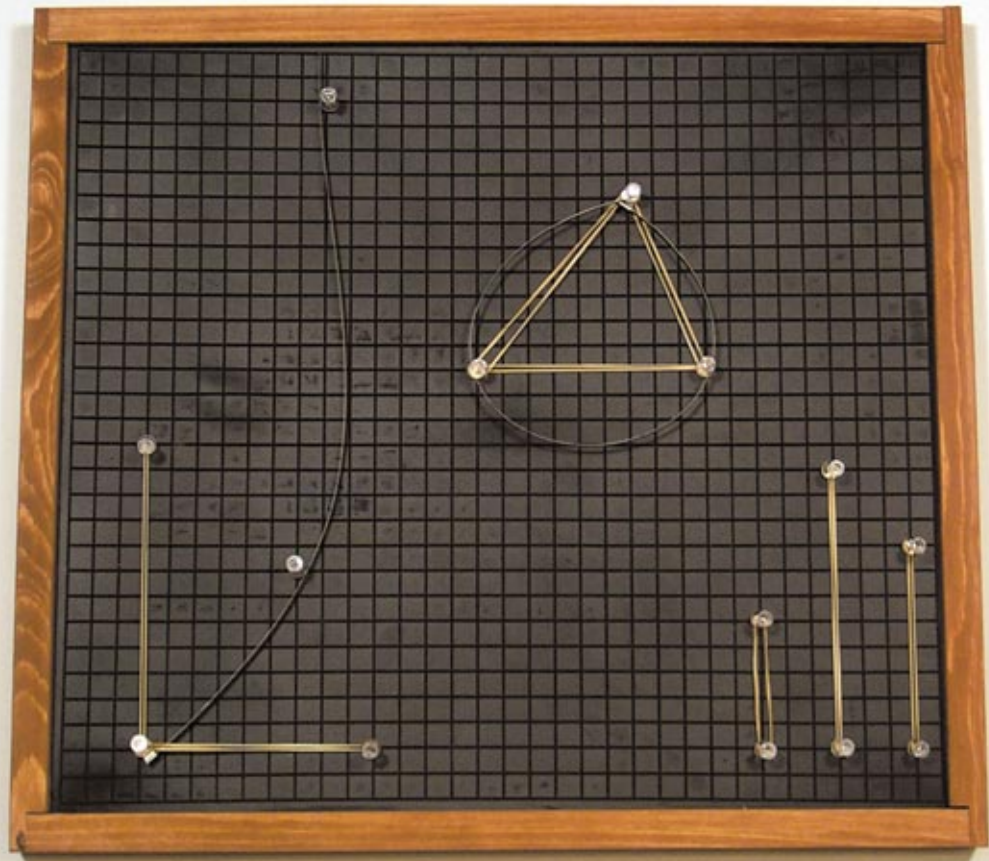


Figure 13: The Graphic Aid for Mathematics. Push pins represent points, including the origin. Rubber bands represent lines, including axes. Note that this tool allows easy graph plotting and exploration, visually and with touch. However, it lacks an easy way to label. It is also not an electronic format, nor can it easily be scanned or converted into an electronic format.

All students in higher grades use calculators for mathematics, including graphing calculators. It allows a student to quickly plot and explore graphs. Desktop tools also exist. Unfortunately, electronic graphing tools are not accessible to the visually impaired. Some specialized tools exist to print graphs for the visually impaired, but they are intended for sighted teachers. Auditory graphs, covered in the next section, are a promising alternative, but current software products do not appear to be sufficiently aligned with what the students are learning.

The major shortcomings of the current approach for making graphs are access and portability. Blind students who want to use graphs will need access to graphing tools, such as the Graphic Aid for Mathematics, and textbooks with tactile graphics. Most of these students will be in mainstream schools, where this specialization is given to a teacher for the visually impaired (TVI). Since the student and the TVI meet infrequently, and since the TVI must spend a large amount of time creating graphs, the student tends to have limited access to graphs. In addition, the graphs the student reads and writes are in a format that is bulky and difficult to preserve. The Graphic Aid for Mathematics, for example, is designed to be reused for several graphs. Alternative electronic formats make it easier for the TVI to create graphs and the blind student to read graphs, but it is currently impossible for the blind student to (directly) draw graphs on a computer.

In summary, visual impairment will be defined functionally, with “blind” students being those who use tactile graphics. The challenges with tactile graphics are in access and portability, with the access part related to production costs. Therefore, a system which reduced production costs, and increased access and portability, would remove some of the barriers of graph use, particularly for mainstream situations.

2.4.4.4 International Differences

In the United States, visually impaired students have the same standards as their sighted peers (including in the upcoming Common Core Standards (CCS)). Thus, visually impaired students are assessed about their knowledge on the same topics, including graphing.

This is not the case for all countries. In Kenya, for example, low vision and sighted

Table 7: A comparison of requirements from Form 1 (9th grade) mathematics, section 19.0.0, COORDINATES AND GRAPHS. See [? ?]. Notice how the blind syllabus emphasizes *exploring* graphs, while the standard syllabus emphasizes *exploring and authoring* graphs.

| Sighted | Blind |
|--|--|
| hline draw and label the complete cartesian plane | identify completely labeled cartesian plane |
| locate and plot points on the cartesian plane | identify points on the cartesian plane |
| | identify appropriate scale from a given graph |
| choose and use appropriate scale for a given data | work out appropriate scale for a given data |
| make a table of values for a given linear relation | make a table of values for a given linear relation |
| use values to draw a linear graph | |
| solve simultaneous linear equations graphically | determine solutions of simultaneous linear equations from tactile graphs |
| draw, read, and interpret [sic] graphs | read and interpret tactile graphs |

students are expected to author and explore graphs, while blind students are only expected to author graphs (See Table 7). My own (unpublished) research in Kenya shows materials that are difficult for visually impaired students to use for authoring graphs, although they may be used for exploration (see Figure 14). The tools available do not allow graph creation, and the standards do not require it. In the United States, all observed graphing software used is virtually inaccessible to blind students in terms of both exploration and authorship for topics in the Common Core Standards (CCS).

2.4.5 Discussion

The exploration of curriculum, graph literacy, and visually impaired students led to several related findings. Curriculum development is a complicated concept, guided by standards. Learning experiences should be generated to meet the standards, and are flexible enough so that different experiences can meet the same standard, and a single experience can apply to several standards. A key component the the experience is teaching the core concept, which also relates to testing accommodations. In order to avoid the “graphing calculator problem,” graph making should allow a more direct plotting of points, with a metaphor closer to that of graph paper than a graphing calculator. Graph literacy research shows that graphs can



Figure 14: Two graphics from Kenya. Note how the figures can be explored but may be difficult to create by a blind student.

be more insightful and efficient than looking at the data or statistics. However, graphicacy is not well-understood in practical terms. In addition, graphs for the visually impaired are difficult to create and are not very portable.

One primary thread in this background is a gap in understanding what students are actually *doing* with graphs. The GPS and CCS require graph literacy of certain tasks, but the steps to complete the tasks are not described. Graphicacy research has little insight on the practicalities of learning graphs. And alternative formats are not as robust as the typical visual graph.

Instead of working from generalizations, graph literacy can also be defined from the bottom-up. By looking at the Common Core Standards (CCS) related to graphing, finding graphing questions that relate to the standards, and understanding the steps used to solve those graphing questions, a list of requirements for graph literacy can be tied to a specific set of standards, such as a grade level. This research would link theoretical and practical graph literacy, and lead to a set of high-level requirements for any graphing technology intended for students learning the standards.

2.5 Conclusion

A few themes emerge from the broad background. First, there are no mainstream technologies that allow visually impaired students to plot graphs on a computer. Research from a number of groups has produced innovative ideas, but it remains unclear to what extent the evaluation results relate to actual graphing in class. Graph literacy itself, while recognized as important, has no grounding in actual graphing problems in education. Finally, given an approach for effective point estimation with auditory graphs, and an understanding of the tasks taken for graphing questions based on standards, it would be possible to evaluate the effectiveness of a system in classroom and testing situations.

CHAPTER III

TECHNOLOGY

The technological objective for this dissertation is to create software for sixth grade blind students that they can use for graphing in math class. That objective required several foundational technologies. A technology for blind students who are graphing will need a novel way to graph, specifically a way for them to easily estimate where points are on a graph, and a way to add points to the graph. Understanding active point estimation interaction required its own software program, as well as a program for training students on the new interaction. Even the computational data structure of a "graph" was not apparent. Modern forms often reduce the structure to final artifacts, such as a picture, or omit important graphing information, like in tables of data such as CSV. Once the data model was established, an additional technology was required to structure the systematic conversion between data sources, graphs, and final artifacts.

This chapter presents the technologies created for this dissertation research in active point estimation in the classroom. The first half describes a data model for graphing, converting the model to other formats. and a combination of the model and converter into a single software library. The chapter then changes focus to a software package designed for a series of lab studies on active point estimation, followed by a longer training program on the task. The final technology is a classroom tool for graphing. The chapter closes with a discussion of the technologies and lessons learned in the development process.

3.1 AGM: Representing Graph Data, Context, and Presentation

As discussed in the Background chapter, several other researchers have explored accessible graphs. Unfortunately, these innovations have not been built with a computational representation that could easily be used in someone else's technology. When the basic data representation is shifting unpredictably between research programs, it is difficult to compare examples between graph software or research studies.

Sonification Lab interest in this problem began with challenges in saving work created in the Sonification Sandbox. Early versions of the tool allowed users to import data points or enter them in manually, with an option to save the data in CSV. It also provided a way to save a MIDI file generated from the auditory graph settings ???. However, there was no way for the user to save the current data-and-sound-settings work for a later time, or to send to someone else. She could, at best, save the CSV and write down her audio settings somewhere else, so that she could later open the CSV in the Sonification Sandbox and adjust the audio settings back to the settings in the previous session.

Looking into the problem deeper, there was not a shared data structure for data and settings. A two-dimensional number array (a table) stored the data values, while the settings were each stored in separate variables. Tracking these variables and the array in the code is tedious, and leads to a high potential for creating bugs when adding to the code base. The conceptual model of graph-and-sound-settings was not being represented in the program. As the lab improved and expanded auditory graphs research, future changes to the Sonification Sandbox would be easier if the data model were a part of the program. This separation was already making it difficult for users to simply save their work.

3.1.1 Storing Graph Data

The solution began with a general graphing model, with a focus on audio and accessibility ???. The Auditory Graph Model (AGM) combines graph settings with graph points lines, and canvases, capturing the graph display window, audio settings, visual settings, and a variety of other components¹. For example, a line could now be set to have a certain color (blue) and a certain timbre (piano), and that setting was a part of a "graph" object, which could be saved, visualized, printed, heard, or emailed to a friend.

Due to early interest in cross-platform compatibility, and interest for the program to act as the graph model for the Java-based Sonification Sandbox, the AGM was programmed in Java and intended for desktop purposes. The AGM uses an event-based system for

¹Since 2006, the "AGM" stood for "Auditory Graph Model". This has recently changed to "Accessible Graph Model" to more fully represent the contribution of the model to several modalities beyond audio alone, and the continuing project emphasis on accessibility.

```

Graph graph = new Graph();
PointsGraphLine line = new PointsGraphLine();
GraphPointElement time, value;
line.addPoint(new GraphPoint(0, -3));
line.addPoint(new GraphPoint(2, 1));
line.addPoint(new GraphPoint(4, 3));
line.addPoint(new GraphPoint(6, 5));
graph.addLine(line);
Formula formula = new Formula("y=-x+1");
FormulaGraphLine line = new FormulaGraphLine(formula);

```

Figure 15: Code for a graph with two lines. The first line has 4 points, along the path of $y = 2 * x - 3$. The second line is simply represented by the formula $y = -x + 1$.

indicating when changes have occurred for a part of the model. Many of the settings cascade: settings on a more specific scope are honored above those of a more general scope². For example, a particular point can be set with a different color than the color settings for the line of point data.

A graph object is the top level of the model. The graph object holds a window, or viewing space, and lines of data. Since the Sonification sandbox used a spreadsheet for entering in data points, the lines were initially composed of a set of points. You will find this dissertation emphasizing single point values. This is useful for collected data but poses a problem for mathematical functions³. For example, a graph of $y=2*x$ may be optimized with the points for display at $-10 \leq x \leq 10$. However, if a person zoomed in at a portion of the display, or shifted to a different part, the display may not show points in places where they should be listed. There are in fact infinitely many points.

To solve this problem, we created two types of lines: point lines and formula lines. Point lines are defined by a series of points. Formula lines are defined by a formula, such as $y = -x + 1$. For formula lines, points can be derived from the formula as needed, so that a display may show display artifacts for any location or resolution needed. Point lines have

²The inspiration was cascading style sheets (CSS), used in web development.

³By ninth grade, students should be familiar with functions and drawing lines on graphs. Before then, most graphs are drawn with specific points.(VERIFY THIS)

two more layers. Point lines consist of points. These hold one or more dimensions of data, such as (x, y) pairs. Point elements are dimensions of the data. They are a identifier-number pair, where the identifier is a dimension-label pair. For point lines, only those points which are predefined and within the window are displayed. For formula lines, the data is derived for display for the specific window. Figure 15 shows example code for initializing the lines.

3.1.2 A Window of Opportunity

As introduced in Chapter ??, context is the graph information that assists the user in understanding the position and value of pieces of the displayed graph. Earlier work in our lab has shown the importance of context (CITE). Context in the AGM is represented with axes and tick marks, and a window. Most of the context representation is straightforward, but the implementation of graphing bounds is discussed here in more detail.

The importance of graph bounds is often overlooked. Consider Figure 16, which shows two graphs of stock prices. The graph provides useful information on the stock trends. Even a novice trader can see a drop in the market lasting about six months, followed by a strong recovery. There are also long periods of relatively little change. The magnitude of changes is lost, however, in the automatic selection of the stock value range shown. Notice how the graph's vertical data bounds are determined by the values in the graph. The vertical ranges in the graph are very different, 1050 to 1500 and 280 to 400; while the dips in the graph look similar, they differ in point magnitude (150 points compared to 60 points; although the percent change is about the same).

The creators of the interactive graph designed the display to hide some information along the x range⁴ of values. They could have decided to include the entire range, from March 4, 1957, to the present. I will call this a *automatic range*, compared to the *selected range* seen in the x values in Figure 16. An automatic range would provide a continuous display of all of the market information available. Earlier trends could be compared with later trends, and the magnitude of the differences would also be preserved in a single graph.

⁴Technically in mathematics, "range" refers to the minimum and maximum y values, and "domain" refers to the minimum and maximum x values. In other contexts "range" refers to the minimum or maximum values of any data set, or the difference between these values. This discussion will use the latter definition, since it applies to both x and y and is sufficient for making the point.

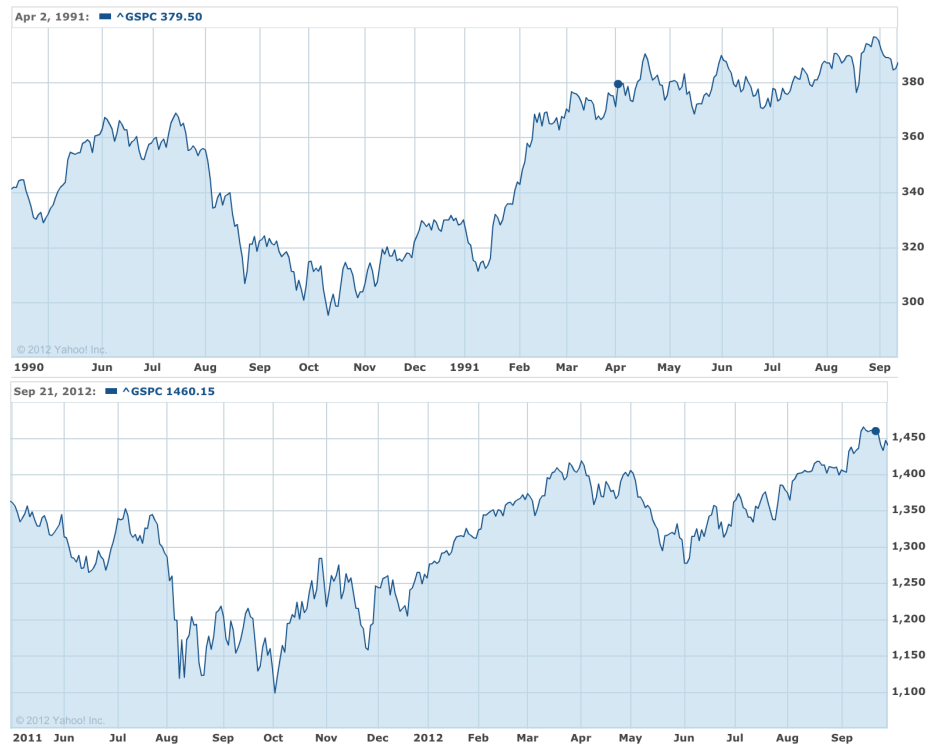


Figure 16: A Yahoo Finance graph of the S&P 500 value in 1991 and 2012. The trend of the drop and recovery is similar, but the magnitude is different (60 compared to 150).

On the other hand, An x range of 55 years would force a larger chart x width for the additional data, or require more data in the same amount of space (or a combination of the two). The charts currently show 18 months of data. For the 55 years currently available on Yahoo finance, the x width would have to be adjusted to around 37 times the current width, or approximately 21 feet! In the other case, the data density would be higher than the number of horizontal pixels, so the program could not render every data point. Some compromises would be required. After the display is rendered, as in Figure 17, it has some interesting trends, but the patterns in the two graphs in Figure 16 have been lost.

While the x values of the display have a selected range, the y values have an automatic range, which cannot be changed by the user. The y value range displayed is always slightly larger than the y value range for the data in the x value range. All of the y data is displayed, and regardless of the magnitude of the data range, the vertical display range is automatically optimized for the most vertical space that can display all of the vertical data values. Therefore, the drop and recovery in Figure 16 in 1991 appear to have the a slightly

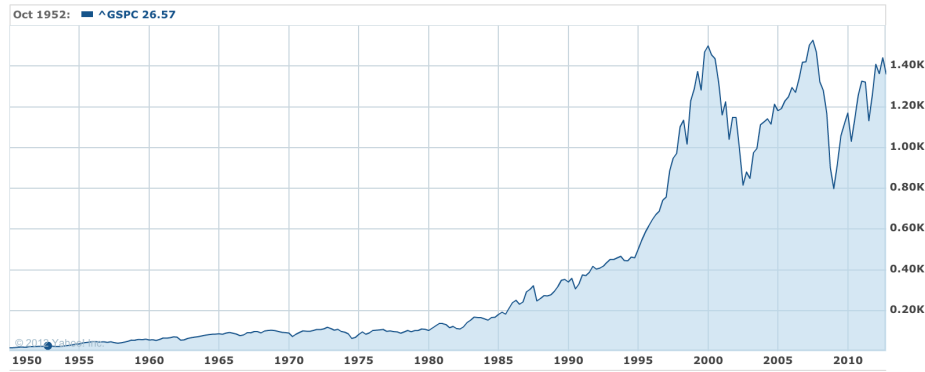


Figure 17: The S&P 500 values, 1955-2012. The overall pattern is visible, but smaller patterns found in Figure 16 can no longer be perceived. Figures from Yahoo Finance.

larger vertical size than the drop and recovery in 2012, even though the magnitude was 30% smaller in 1991. An adjustment of the vertical values to a common base, such as zero, or to have the same range in each graph, would make the magnitude differences more obvious.

In stocks the trends are often more important than the actual values. As you will see in Phase 2, and may already be aware, point estimation is an inescapable component to graphing in K-12 mathematics education. Unlike the (mostly) gradual vertical changes in stock prices), functions can have dramatic changes over small x values. Finally, many functions, such as $y = x$, have an infinite data range for x and y , so cannot be displayed in an automatic range.

The words "automatic" and "selected" have an intentional suggestion of shifting responsibility of window selection to the graph's author. It is no coincidence that Yahoo Finance uses automatic y ranges: with a fixed starting x range of the past day (for example), and an automatic y range, Yahoo does not require human creation or screening of the data display. Graphs can be "automatically" created in real time with minimal cost. However, the interpretation of the graph may not be as simple; the graph's user has a larger responsibility to interpret the graph himself. A well-trained trader may be suitable to the task, but important graphs for the general public and graphs intended for education should be designed to optimally display the communicative intent. This x and y range is "selected" by the graph's author after thoughtful consideration of the problem. For example, a student learns how he can change the size of his paper graphs and labels of his grid lines to fit the important data

```
// initialize graph above...
Window window = graph.getWindow();
window.setBounds(new RectangleBounds(-10, 10, -10, 10));
window.addContext(new AxisContext(0)); // x-axis
window.addContext(new AxisContext(1)); // y-axis
window.addContext(new TickMarkContext(0, 0, 1));
window.addContext(new TickMarkContext(1, 0, 1));
```

Figure 18: Setting the Window properties on a graph.

points for a homework question. This is an important component to graph creation and graph interpretation, and should be a part of any graph learning system, including those for visually impaired students.

Automatic range and selected range can operate in the same system. For example, in many K-12 graphs, a range of -10 to 10 for x and y is designed to be appropriate. These may be default values in a display setup, along with information or autofill options on the data x and y range for the currently targeted region.

In the AGM, a "window" is the viewable space, or data range, that will be displayed to the user. A window is composed of bounds: the limits of a display for all dimensions being explored. The window also holds and context elements such as axes and tick marks. An example of setting up context in the model is available in Figure 18. First, the bounds are set to -10 to 10 on x and y. Next, axes are set for x and y, and tick marks are set, 1 for each unit. Graphs displays such as in visual and auditory formats use this information to present indicators of tick marks, axes, labels, and graph bounds.

3.1.3 Generic Settings for Displays

The AGM also contains settings for various display renderings. These general options are intended to support a wide range of converters (discussed further in Section 3.2 about the AGFC). When rendering settings are incomplete or completely missing, the display can revert to default settings available as part of the AGM. For example, an unspecified line color might default to red for the first line, and blue for the second line. Figure 19 shows a code example of rendering an auditory graph with MIDI output.


```
// initialize graph above...
MIDIRendering rendering = RenderingFactory.createDefaultMIDIRendering();
rendering.getTime().setTime(5000); // 5 seconds playback.
g.replaceRenderingOfSameType(rendering);
// Treat pitch as a mapping (keep pan and volume fixed and at default values)
GraphFactory.setPitchIdentifier(rendering.getPitch(), new Identifier(1));
```

Figure 19: Setting MIDI Rendering settings for a graph. When renderings are incomplete or completely missing, the display reverts to default settings.

3.1.4 Limitations

The AGM provides a model for capturing data for graph displays, including the data itself, context information, and settings for the eventual rendering. This is a useful first step, but requires a conversion to a display for people and most other programs. That work will be discussed in the next section.

This model is useful, but is locked into Java. Recent work has explored the use of a model that could be deployed in a native flat file format, with the capability of being readable in several programming languages. That possibility is becoming a reality in the ChartML project. This partnership between the Sonification Lab and SAS Inc. is looking to model major charts in education in an XML schema. This schema could then be converted to Java through automated processes such as JAXB. The AGM is providing an example of a useful structure. Once completed, the more flexible ChartML will probably replace the AGM.

3.2 AGFC: Converting between the Model and Other Formats

The AGM provides a representation of graphs that only a computer or a programmer with the right access can interpret. The AGM requires a separate tool to render the information perceptible to people. In many cases, other file types must also be converted into an AGM format. The developed two-way converter is known as the Accessible Graph Format Converter (AGFC). Taken together, the AGM and AGFC constitute the Accessible Graphing Engine (AGE), a powerful tool for data input, modeling, display, and output. This section

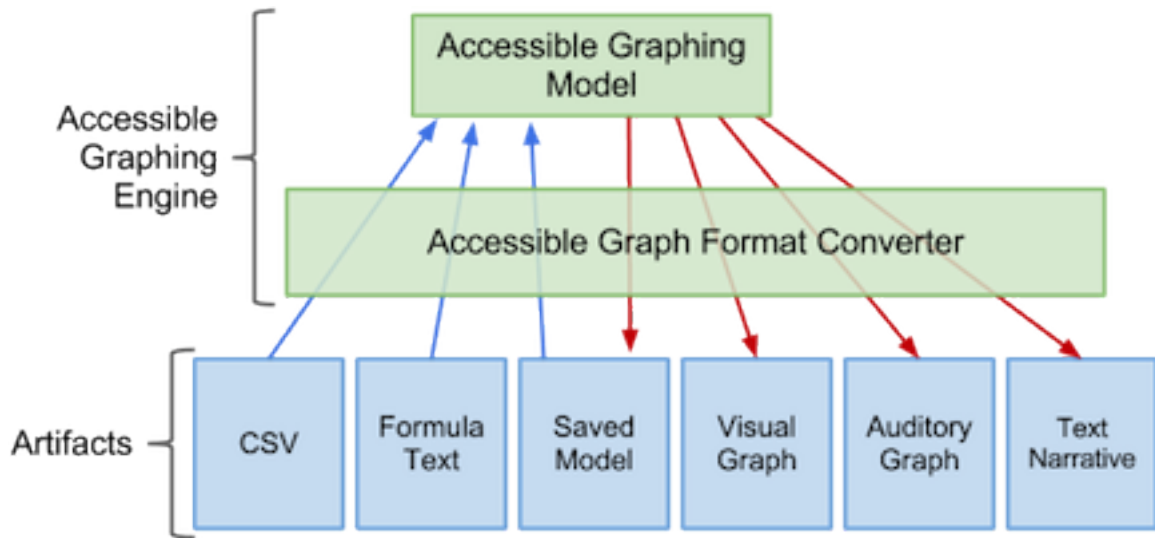


Figure 20: An overview of the Accessible Graphing Engine (AGE). Artifacts can be converted to an AGM model by the AGFC (blue lines), a process called abstraction. AGM models can be converted by the AGFC into artifacts (red lines), a process called reification. Some artifacts can be converted both ways such as the Saved Model.

will discuss the AGFC, while the next section describes the AGE.

Members of the Sonification lab found a stronger model in the Sonification Sandbox helpful, but it required several hooks into the model for someone to be able to modify, view, and save their work. There are in fact many different potential sources of data for the AGM, and many potential targets for the AGM. A reliable and common process for importing and exporting AGM data was necessary. Initially, the scope of the AGM expanded to include this translation, but it became increasingly apparent that the missions of the new task was very different than the AGM, and a separate project, the AGFC, was created. This would allow us to more deeply develop each technology, and to switch out model and translation components more easily (ChartML for example).

There are a few new terms that describe the translation process. An *artifact* is information that will be translated into the AGM, or data that has been converted from the AGM to some other format. Since an artifact is often a real "thing", more than a graph concept, such as an actual visible graph, this was called a reification⁵. The graph model is

⁵The definition of "reification" is to "make concrete." In Computer Science, reification is a model that makes a concept concrete; from this perspective, the graph model is a reification. Further consideration

Table 8: An overview of the AGFC abstractors and reifiers. There are five abstractors and 14 reifiers, four of which can do both.

| Translators | Name | Description |
|-------------|----------------|---|
| A & R | CellGrid | An internal format. |
| A & R | CSV | Comma Separated Values, a spreadsheet format. |
| A | Formula | A mathematical representation of a line. |
| R | GraphML | An XML schema for describing graphs. |
| R | Image | A static image creator (PNG and JPG formats). |
| R | JFreeChart | An interactive graph. |
| R | JFuguePattern | A MIDI-based auditory graph. |
| R | MDEAudioGraph | A waveform-based auditory graph. |
| R | MDEVisualGraph | A static image. |
| A & R | Serializer | A file format for saving all AGM information as Java. |
| R | Speech | A spoken description of the graph. |
| R | Text | A written description of the graph. |
| R | Wav | A waveform of the MIDI-based auditory graph. |
| R | XML | An XML format of the Serialized file format. |
| A & R | YML | A markup language similar to XML. |

an abstraction. Therefore, an *abstractor* converts artifacts into graph models. For example, importing a CSV file into the Sonification Sandbox is one instance of abstraction. A *reifier* converts graph models into artifacts. For example, saving a MIDI file generated from a graph in the Sonification Sandbox is an instance of reification. A particular artifact may have an abstractor, a reifier, or both (for example, a CSV file can be converted to the model, and can be created from the model). Figure 20 shows an overview of artifacts, the AGFC, the AGM, and the AGE.

Here’s an example that walks through the entire flow through the AGFC. Danny is interested in exploring the daily and overall changes in temperatures, as they compare between various U.S. cities. Danny has high temperature data for the first 20 days of October for Atlanta and Boston in a spreadsheet. He saves the spreadsheet data as a CSV, a simple format that uses commas and line breaks to indicate difference columns and rows in the data. Danny uses an AGFC tool to import the data. Once imported, this data is in AGM format. Danny then views a picture of the data. He modifies the picture to have

of naming earlier on could have avoided this unfortunate contradiction. I suggest considering the abstraction/reification from the non-computing perspective, from the perspective of the end user. An abstraction takes real graph data and makes it available for exploration. A reification creates real, perceptible graph displays.

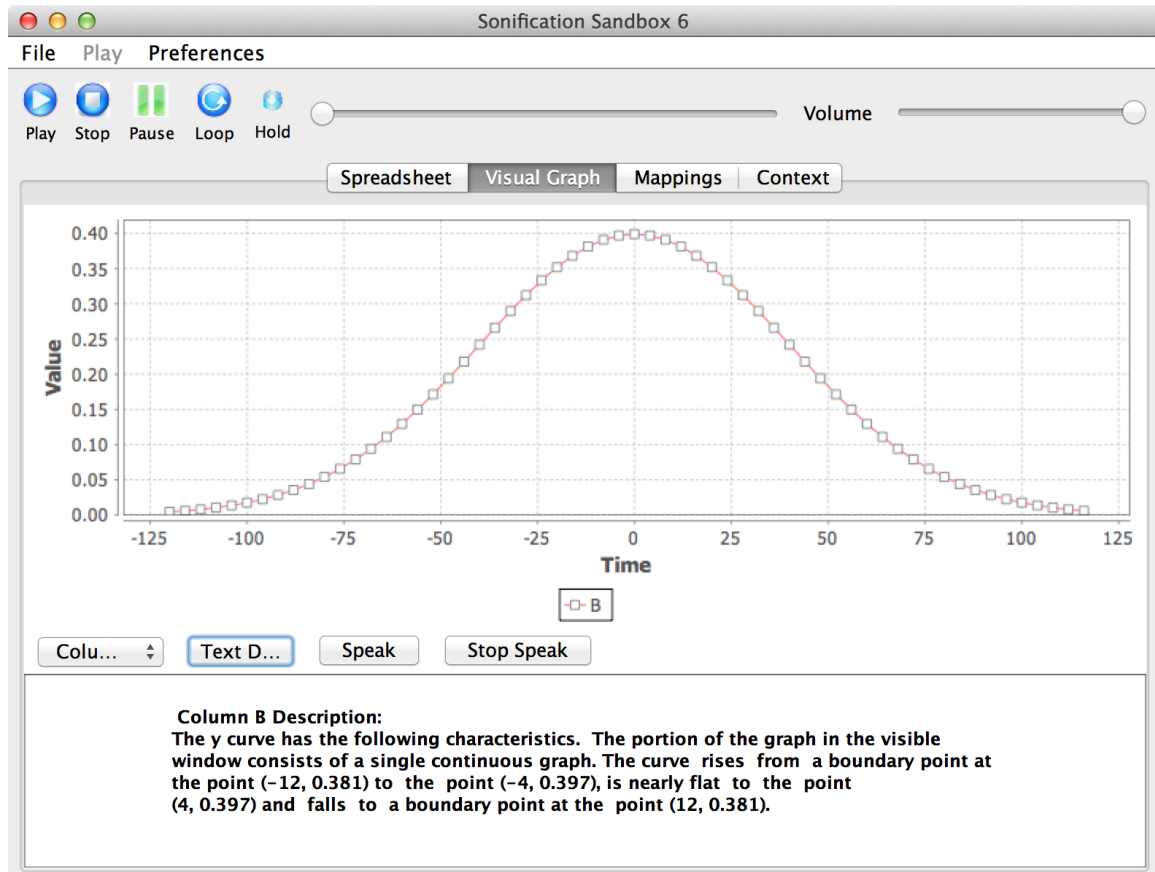


Figure 21: The Sonification Sandbox software features opportunities to explore data in multiple formats simultaneously. This screenshot shows the user interface that features a visual graph, an auditory graph, and text description of the graph

the y-range (window) include 0 degrees, and changes the colors of the Atlanta and Boston temperature lines. Danny then saves the picture as a PNG image on his desktop, and shares it with a coworker over email. Danny saves the AGM file for use at a later time.

This story covered all pieces of the AGE, including artifacts, translation, and the model. There were four artifacts in this story. The CSV file could be imported into the AGM (through abstraction). An image file was saved, which is an output artifact (through reification). In addition, Danny used an interactive graphic to view the picture while making changes; this is also a reification. Finally, Danny saved the AGM file for later use; this is an artifact in a format that can be completely be translated back into the AGM for later use. This file save was a reification, which may lead to an abstraction later on, by either Danny or his coworker.

```

// Get the saved AGM file .
SerializationAbstractor abstractor = new SerializtionAbstractor ();
abstractor.setFile(new File("SavedGraph.agm"));
abstractor.abstractAsUnthreaded ();
Graph graph = abstractor.getGraph ();
// Save as an image .
String imageFileName = "./GraphPicture.png";
ImageReifier reifier = new ImageReifier(graph);
reifier.reifyAsUnthreaded ();
ImageReifier.writeGraphToFile(reifier.getImage(), imageFileName,
    ImageRendering.image_type.PNG);

```

Figure 22: Code for importing a saved graph file (abstraction), then exporting an image (reification).

Note that this story did not give a hint at the specific tool Danny was using. The AGE is a software library for end-user software. Danny could be using the Sonification Sandbox, a web-based alternative, or a completely separate program⁶. Part of this flexibility stems from a common model, the AGE, and part stems from the conversion tools found in the AGFC. Many abstractors and reifiers are provided as part of the core AGFC. Table 8 provides an overview.

The translation architecture has a few major components. First, every translator object is either an Abstractor or a Reifier. An abstractor takes an argument which captures the input artifact, and outputs a graph. A reifier takes in a graph and an output target, and sets up the output at the target. As mentioned before, Abstractors and Reifiers are Translators, and have a lot in common. Translators must be triggered to begin. The translation can be threaded or unthreaded (in other words, it can be a background task, behind other software activities). The translator must provide its basic status, such as ready, translating, or done. The translator can also provide other information, such as the approximate completion percent. This can then be shown to the user, in user interface widgets such as progress bars.

⁶The Sonification Sandbox code has become much simpler. Since it no longer has model or conversion functions, its major components make options visible to the user.

It may be apparent in the example above that it is common to convert to a lot of reified formats from a single model. During the process of understanding the data or saving it for later use, the best format can be used. To facilitate this process, each for the translations can be run threaded or unthreaded. In user interfaces, threading allows user actions during long tasks. Threading in this case also allows the presentation of multiple formats of the data at the same time. This combination of presented artifacts often becomes an artifact in itself. For example, imagine a visual graph of city temperatures, with an auditory player also available. When the user presses "play" for the auditory graph, they can follow the data with both their eyes and ears. The most salient features in each modality are present to the user, potentially creating a graph that is more informative than the original graph. It is also, of course, potentially more accessible. This potential can be found in the Sonification Sandbox software, also seen in Figure 21. Multiple formats can also be found in other accessible software, such as MathTrax and the Audio Graphing Calculator, although a comprehensive model, access to their model, and conversions to other formats is more limited⁷.

The multiple format conversion can be brought to an extreme case. The sample program "Mass" presents sample data sets, being dynamically generated in every reifier available in the AGFC. Simple data sets can be rendered in all formats in a blink of an eye, whether it is a speech file, an XML file, or a picture (see Figure 23). The Mass UI itself is a simple example of the power available; it is not designed for real graph authorship or exploration. There are many use cases where any of these panels, taken alone or in groups of two or three, could be used for understanding the data points and trends.

3.3 AGE: Combining Model and Conversion

The Accessible Graphing Engine (AGE) is simply a combination of the model and the converters into a single programming library. It is rhetorically simpler to discuss the model and conversion as part of a single tool, and with parts that are not specific to the current release. This section will discuss some issues with using the AGE, and propose future work

⁷Through a partnership with NASA and Benetech, the AGFC leverages certain conversion tools found in MathTrax's Math Description Engine, particularly the text description.

in this area.

The previous section presented several examples of how the AGE can be used in software. A final example is on websites. Although the AGE requires Java, it can be run completely server-side, so that web visitors do not need Java to get feedback. One example is the Web Sandbox, visible in Figure 24 and <http://sonify.psych.gatech.edu/graph>.

3.3.1 Discussion

The AGE provides a powerful tool for creating and storing graphs. This model forms the basis for converting between different formats, such as visual and auditory graphs. The model makes it easier to create consistent results in several formats, an important achievement in accessible graphs research.

One major limitation to the current AGE is that the AGM is not very robust to change. New types of diagrams or data formats will require a lot of time to integrate into a Java-based model. During development, simple improvements such as supporting Java Beans throughout the model have helped, but, fundamentally, the model should probably move out of Java. In addition, while Java is relatively portable, the Java requirement does limit the potential on newer hardware such as mobile and tablet, and access to the model from other programming languages.

Other data models, particularly XML schemas, are more appropriate for language-independent representations. A suitable XML schema can represent the same core model as that found in the AGM. In addition, with tools such as JAXB, the (astonishing!) automatic XML-to-Java creation of model storage, accessors, and mutators, there is a much broader potential for portability and inter-language support. In addition, schema validation and XML validation are broadly supported. Current efforts to create an XML schema are available in the ChartML project, a partnership between the Sonification Lab and the SAS Accessibility Team.

The interchangeability of the model from the translators is an important aspect of the AGE. As the architecture of the model or the translators changes, there should be a clean break between them so that the impact of the changes is minimal. Defining the boundaries,

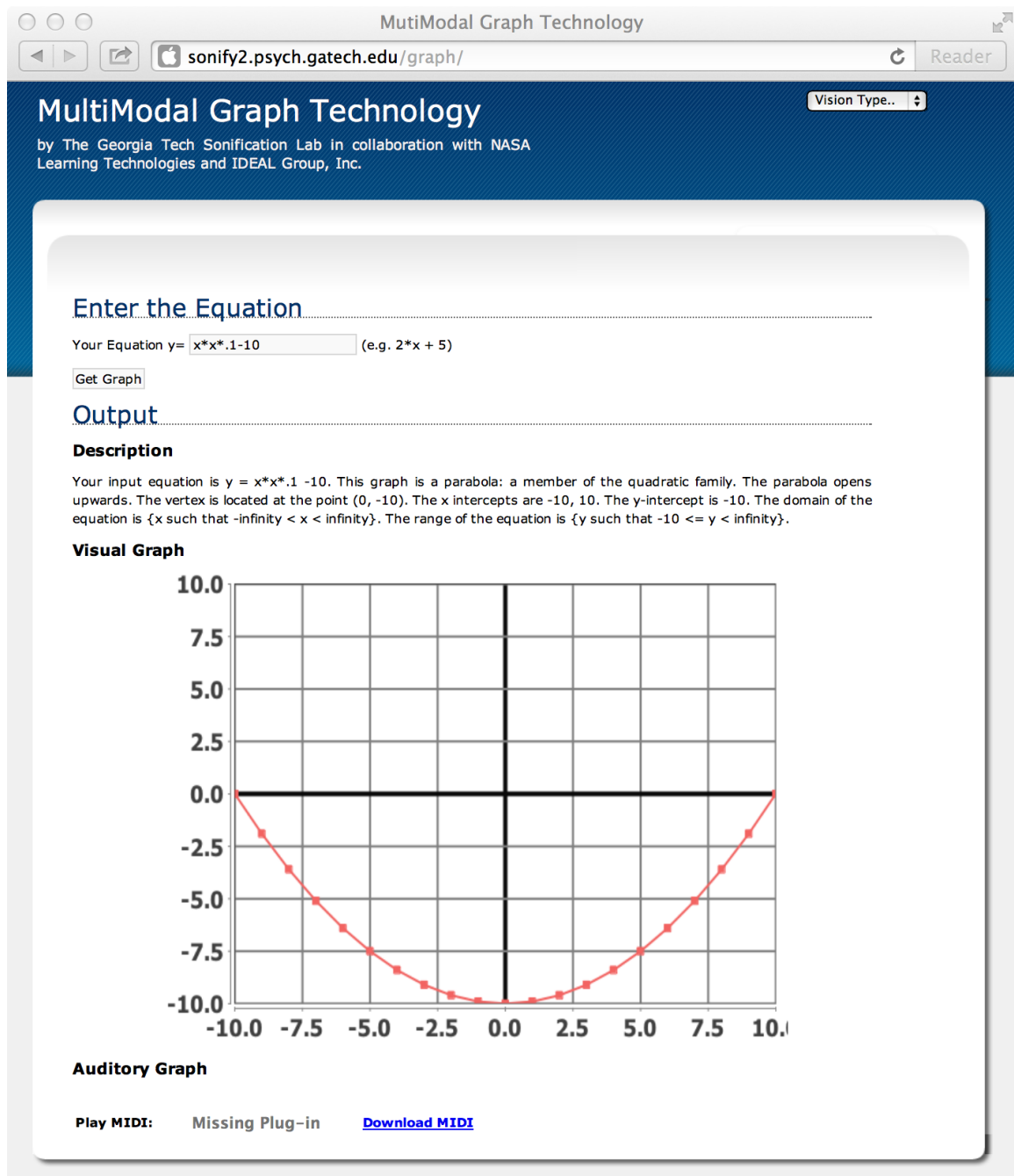


Figure 24: A simplified, web-based version of the Sonification Sandbox. Enter in a formula to see a visual graph, and auditory graph, and a text description. As of 2012, this is a living website available at <http://sonify.psych.gatech.edu/graph>.

what is AGM (model), AGFC (translators), and AGE (both the AGM and AGFC in a full pipeline resource), facilitates reconstruction of the AGE with new parts (e.g. a new AGE, depending on a model based on ChartML).

Our discussion of the AGE has focused on problems, architecture, and software. It has not addressed accessibility, which may be surprising, considering that the first word of these three technologies is "Accessible". Successful development of accessible tools is a process as old as the product. It turns out that much of the accessibility challenge has to do with a broad view of how users will interact with your tool. In a sense, the AGE does not need to be accessible, since it is a tool for programmers and not end users. In another sense, it is critical for these tools to be accessible, since they must deliver the core artifacts for the dependent tools. The AGE has some accessibility by using the Java Accessibility API, a software library for making all objects accessible, and by integrating display components and options that are known to change the level of accessibility. The W3C, among others, has pointed out that by creating accessible websites, developers also create a better core architecture, which leads to efficiency and portability improvements as well (CITE W3C business case for accessibility). I suggests that the causality is reversed: having a solid architecture in the first place, in part by considering accessibility, leads to a software library that can be more easily adapted for accessibility use cases, along with improving efficiency and portability.

3.3.2 Future Work

There are three avenues of further development in the AGE. Substantial improvements are possible in the translators. Some of the converters could use a few iterations of minor development. In other cases, there is a large potential on developing a new translator, particularly abstractors. One particularly interesting area is computer-driven interpretation of graphs, particularly through computer vision (CITE a few). One outcome would be the possibility of retrofitting images with the AGE, offering additional accessibility to text that was not originally prepared to be accessible. Second, as mentioned above, a new, non-Java model of the graph would make the model more robust and flexible. Third, as suggested

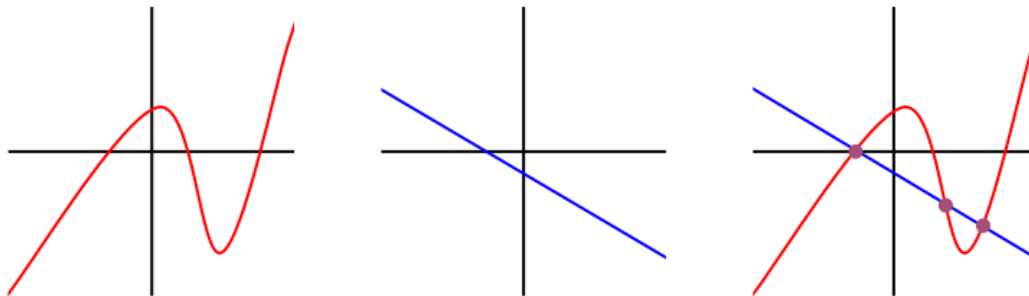


Figure 25: The leftmost image and the center image are two separate visual graphs. In a software program, these could represent graph models that can be manipulated. Dragging them together could combine the data, making a graph like that in the rightmost image. It may also be possible to separate the data apart again.

above, the AGE would benefit from the addition of a tool that can combine graphs. This would make the model a potential nexus of data analysis, providing new avenues for merging and separating data.

One final consideration is a more ambitious development of the graph model. In the language of Computer Science, the AGM is acting as a "read-only" file for the reifiers. The reifiers can act independently because the underlying data is remaining the same. This is not the case for abstractors. Since abstractors convert from other formats to a graph, they are "writing" to the graph. This is more complicated to manage from a model stability perspective, and at this point, a graph has simply one source. It can be edited after creation, but there is no option to, for example, input data into a single line with both a formula and spreadsheet. This limitation is acceptable for the current goal of creating graphs for K-12 education, but there are interesting possibilities for melding data together in future iterations. With a strong model, for example, it may be possible to meld together two graphs. Figure ?? presents one possibility.

The AGE is a starting point for graph data structure and conversion. The next sections cover research-oriented projects which expand beyond the scope of the AGE. These focus more on the student users and the specific challenges of active point estimation.

3.4 AudioFitts: Discovering Active Point Estimation

The common technique for auditory graphs is playing back a static sonification, where x is represented by time and y is represented by pitch. For early graph education, however, the auditory graphs status quo is insufficient for data exploration. An example will illustrate the major challenges.

3.4.1 A Blind Student Learning to Graph

Cindy is a bright sixth grader who comes into math class every day with a positive attitude. With her independence and charm, you would hardly guess that Cindy became blind 3 years ago. Her math teacher Joe tries to keep up with her appetite for learning, even though it is a challenge to find the money, time, software, infrastructure, and expertise to use accessible materials in a public school. In today's class, Joe is reviewing some work from the previous week on plotting on a coordinate plane. Joe has prepared a short worksheet with four sample exam questions, and hands it out at the beginning of class. Joe sees another student, Vivian, solving the first problem: Plot $(-2, 3)$ on the graph. Following her training, Vivian finds the origin, moves her pencil left two tick marks (the first one is labeled "-1", so two tick marks are correct), and then walked her pencil up three tick marks (the first one crossed was labeled "1"). Then she drew a fat dot and went to the next question.

Joe turns his attention to Cindy. Today she is using the Sonification Sandbox, a tool known to the readers of this narrative but new to Joe. Cindy said her vision teacher had trained her on how to use the tool, and just needed a description of each graphing question in an email before class, which Joe had sent to her. Cindy uses her refreshable braille display to read the first question, waiting in an email on her phone. She understands the question, but graphing and the Sandbox are both new to her, so she takes deliberate steps in making the graph. First, she opens the Sandbox. After the program has loaded, Cindy navigates to the spreadsheet tab. She enters in "-2" in cell A1, then "3" in cell B1. Cindy then double-checks her work in the text boxes, listens to the short auditory graph⁸, and

⁸In the current Sonification Sandbox, with default settings, this would sound like a single piano note, played for 5 seconds.

changes to the visual graph for her teacher Joe to look at. Joe tells that she has the answer correct. Cindy moves on to the next problem, but Joe is troubled: has Cindy really learned how to graph?

Vivian's behavior to answer the question involved identifying the origin, finding the right x tick mark based on the labels, finding the right y tick mark based on the labels, and placing a dot where those tick marks would cross. Her interactions involved her eyes and her hands. Cindy's behavior to answer the question involved finding the spreadsheet, finding the top left cell, entering in the x value, moving to the cell to the right, entering in the y value, and going to the graph tab. Both, in the end, create a graph, and many of us have made graphs both ways. Vivian's method tends to be taught before Cindy's method⁹.

While a single step in each task is relatively simple, Vivian's task demands more graph understanding from the student. Identifying the origin requires the skill of finding where the two axes cross. Finding tick marks and labels requires finding the correct axis, matching labels to tick marks, and interpolating between tick marks. Placing the dot requires hand-eye-coordination along two axes. These steps force the student to learn several things important to algebra. For example: compared to each other, variables can change in different ways (independent by having horizontal and vertical movement); numbers exist between whole numbers (e.g. between tick marks); negative numbers are closer to zero than they are to any positive numbers (physical distance); relationships between variables can be predictable (a regular pattern in data points or lines). Cindy's task simply involves moving between tabs and parroting data values. Such skills are useful only after a student has learned the basics of graphing, reinforced through practice like Vivian's behavior. Once the sixth grade class understands how to plot, Joe can show them graphing calculators, and the students can focus on more advanced topics.

If Cindy wants to learn what her peers are learning, she has to act like her peers. To

⁹Perhaps the later introduction of computers and calculators in math class plays a role. Early graphing calculators probably did not have the sophistication to plot directly on the screen. Of course it is possible now to create a virtual canvas where students could use their finger or mouse to plot points. However, as I argue here, the motivation of starting with Vivian's behavior in school has probably more to do with the student having a larger active role in graphing. Later, when plotting is mastered, that process can be automated so the student can tackle more cognitive graphing challenges.

Table 9: Graphing behaviors for Vivian, a sighted student using paper, Cindy, a blind student using the Sonification Sandbox, and proposed behavior for a blind student.

| Task | Vivan | Cindy | Proposed |
|-------------------|---|--|---|
| Find the origin | Look for where two axes cross, point the pencil there. | Not available. | Listen for where two axes cross, move the pointer there. |
| Find a value (1D) | Move along the correct axis, looking for tick marks and their labels, until the right one is found. | Not available. | Move along the correct axis, listening for tick marks and their labels, until the right one is found. |
| Plot a point | Mark where the values cross with a pencil mark. | Type the x, y pair into the spreadsheet. | Mark where the values cross with a virtual dot. |

the extend possible, Cindy should be empowered to find the origin, find a specific point, and plot points the same way her peers are completing the task. A couple pieces may be impossible, since Cindy is blind. But many parts of the task are well within Cindy’s capabilities. The proposed behavior for a student using a new system is in Table 9.

Consider this challenge from a more theoretical view of educational technology. (Greeno and Collins, CITE that education technology paper behavior/cognitive/social) give three frameworks for understanding education: behavioral, cognitive, and social. A student can be judged to be like her peers based on her actions (or more narrowly, her test scores). This same student could be judged based on how she answers open-ended questions on graphing, for example. The student could also be evaluated based on whether she does graphing like her peers, or professionals who use graphs. Cindy should be able to learn how to graph like other students in Joe’s class. This will enable certain contributions of the graph medium, such as spacially perceiving distances between number values and relationships between variables. This could help Cindy understand algebra in a way similar to her peers. Finally, Cindy is doing authentic graphing, adapted to bypass her vision impairment. Cindy is no longer disabled in her mathematics education, and an enormous handicap in her professional opportunities has been removed.

As you will see in Phase 2, navigating a graph is a critical component of *every* graphing standard in Common Core Standards (CCS) sixth grade mathematics. The fundamental

component that is required to enable such navigation is an interactive movement along a single dimension, with constant user input and graph feedback. The activity embodied in this movement is Active Point Estimation (APE). A system using APE for graphing requires:

1. interactions for movement along each dimension that are similar to canvas-based plotting by sighted users,
2. perceptible points, tick marks, and other graph phenomena as they near the point of observation,
3. perceptible text where specified,
4. fine-grained navigation, and
5. a way to add points and text.

The Sonification Sandbox supports items 2 and 5, but falls short particularly when comparing behavior. While the AGE provides a useful backend, it does not appear to be sufficient for supporting live interaction with the canvas. Other tools meet some of these requirements, but not others.

The next three tools explore APE. The first, AudioFitts, establishes the basics of meeting all five criteria with a mouse or keyboard and an auditory feedback. The Navy game trains students on how to use APE quickly. The final program is Graph and Number line Interaction and Exploration system (GNIE), software designed for using auditory graphs in classroom situations.

3.4.2 Design

AudioFitts tests user speed and accuracy in finding targets. Depending on the experiment, users can use a keyboard or mouse for input, and a computer monitor or headphones for system output. It's named is based on Fitts' Law, the famous human factor's rule that essentially means a larger, closer object can be reached more quickly than a smaller object or one that is further away. Naturally the research for this dissertation focused on audio

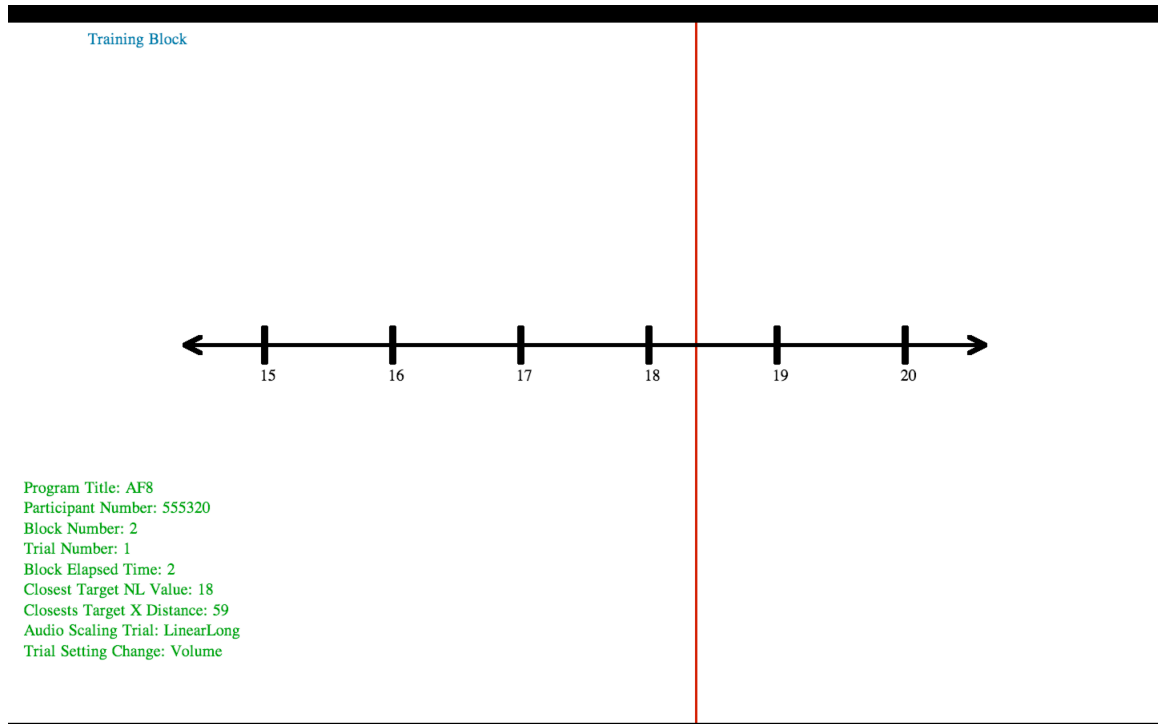


Figure 26: Visuals from an AudioFitts Experiment. The number line has tick marks, which can be accessed by moving the mouse left or right.

components, and AudioFitts was designed for studies in psychophysics. An experimenter specifies his requirements for an experiment, such as when to show visuals or how the pitch should change. After the experiment, the system has a completed log of user speed and accuracy for each trial. Figure 26 shows a screenshot of the program.

Each trial was a simple graphing step, where the user was asked to find the only point or a point with a particular label, and indicate where it is. The user could use the keyboard arrow keys or mouse movement for left-right navigation, and the spacebar or mouse click for selection (the input devices enabled and specific selection mechanism depended on the particular trial conditions). The user could see or hear a one-dimensional horizontal line with one or more tick marks and could move along the line. The user was asked to complete a similar selection task for each trial within a block. After a specified period of time, the user would finish a block of trials¹⁰. Depending on the goals of the experiment, each block

¹⁰In a few cases cases, particularly for training, users had to complete a certain number of trials, sometimes a certain number of correct trials.

was slightly different than other blocks. For example, the width of the target, whether visuals or audio or both are available, and how pitch mapped to distance to the target were a few variables explored in the experiments. Chapter 4 on Phase 1 of the research gives more details of four experiments conducted with AudioFitts.

The core code was in Java, which allowed some interoperability with the AGE and other libraries. The display was rendered with Processing (visuals) and Minim (audio). Each movement would trigger a reevaluation of the visual and auditory display, based on a check of the specific settings for the experiment. The possible variables are listed in Table 10.

AudioFitts logs data to a CSV file¹¹. Before the experiment, the experimenter sets the particular variables that should be logged. At the beginning of each experiment session, a file name is generated based on experimenter input (e.g. a participant number) or the system uses a random file name. After each trial, a line of comma-delineated text is added to the file, indicating the states of each variable being logged. Logging is also available for every display frame¹² or block, but in practice the trial-level information was the most valuable.

AudioFitts creates a data-driven method for finding how to do APE with audio feedback. The five criteria proposed earlier are potentially solved.

1. AudioFitts evaluates a method that works along one dimension. Expanding this approach to two dimensions will not be too difficult, since students are asked to consider one dimension at a time.
2. AudioFitts presents points and tick marks in various audio display formats. Not only will participants be able to perceive these graph pieces, the result of AudioFitts studies will be empirical evidence of the display techniques that makes the graph pieces the most perceptable.

¹¹”CSV” stands for ”comma separated values.” It is a file format designed for tabular data. Each line of text is like a row in a table, with each column divided by a comma.

¹²Java Processing displays ”frames” to the screen. After each iteration, a frame is posted, and then the system processes the next frame. This occurs several (10-100) times a second.

Table 10: Variables available for manipulation in AudioFitts. Most of these variables are described further in the specific experiments in Phase 1 (Chapter 4).

| Name | Options | Description |
|---------------------|---|--|
| Allow Hits | No, Yes | Whether a target hit should trigger the next trial. |
| Allow Misses | No, Yes | Whether a target miss should trigger the next trial. |
| Applet Width | Any whole number | The width of the display space, in pixels. |
| Audio Scaling | Boolean, LinearLong, LinearShort, Logarithmic | How the distance scaled, before mapping to a sound. |
| Blackout Percent | Any decimal between 0 and 1, inclusive. | The percent of tick marks that do not have labels. |
| Block Sets | Any whole number | The number of block groups; run the same blocks for more data. |
| Block Time Length | Any number greater than 0. | The length of a block, in seconds. |
| Interactor | Mouse, Keyboard | The allowed interaction device. |
| Maximum Note | Any whole number 0-127 | The maximum pitch. |
| Minimum Note | Any whole number 0-127 | The minimum pitch. |
| Number Line Target | Any whole number | The tick mark value users are trying to mark. |
| Number Line Width | Any whole number | The width of the number line, in pixels. |
| Sense | Audio, Visuals, Both | The graph display modalities. |
| Show Target | No, Yes | Whether the target tick mark should have a spoken text. |
| Sound Properties | Pan, Pitch, Rate, Volume | The sound property to be manipulated. |
| Target Peak | Flat, Peak | Whether the sound continues distance information when inside the target. |
| Target Width | Any whole number | The width of each target tick mark, in pixels. |
| Text Reading | Automatic, On Demand | Whether to say the label for a tick mark when reaching the point, or when the user requests a reading. |
| Tick Border to Edge | Any whole number | The distance between the leftmost (or rightmost) tick mark and the end of the number line, in pixels. |
| Tick Mark Count | Any whole number | The number of tick marks on the number line. |

3. Text is available through speech.
4. Through the use of keyboard and mouse navigation. users can move accross a virtual canvas toward points and tick marks. They have the capability to narrow their target to a few pixels.
5. Users may indicate where they would add points in a real application by selecting the pixel that is their current frame of reference.

Further explanation of how AudioFitts was used in this dissertation is available throughout Chapter ??.

3.5 Navy: Training in Active Point Estimation

In AudioFitts, we found that basic training in AudioFitts was fairly quick. Most users, regardless of their vision impairment, could use a mouse or a keyboard to find targets with audio or visual Active Point Estimation (APE). We also found, however, that while the accuracy was high, the task was slower with audio feedback when compared to visual feedback. In addition, we wanted to encourage students to learn this new method in a fun and inviting environment. In Spring 2012, we developed a game, called Navy, that would teach students how to find targets by having them blow things up.

The Navy game was inspired by the classic game Battleship. In the original Battleship game, players take turns guessing at the two-dimensional position of an opponents ships. Navy differs in a number of ways. First, interaction is one-dimensional, horizontal like AudioFitts. The user moves her mouse left and right to her target, while listening to ships and markers (points and tick marks). When the user finds a ship, she must report the nearby marker by typing in the appropriate number (0-9). Second, the ships move during the game, from left-to-right or right-to-left, and cross over labeled markers (identical to labeled tick marks). This makes the game more challenging, but is not realistic to how graph points stay at one position. That said, the game requires players to understand their position and the position of the ships, so that players can report the ships' locations. Third, there is one player who is trying to blow up ships; the computer keeps sending out more

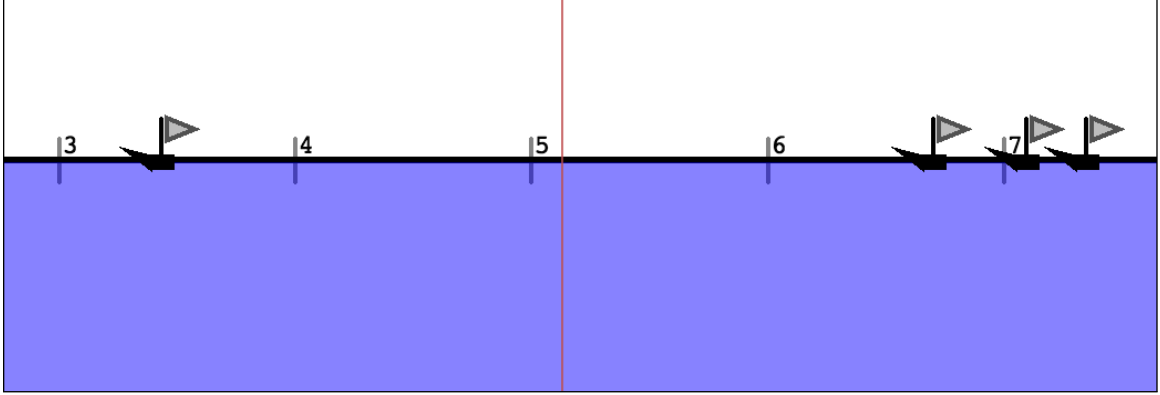


Figure 27: The Navy game used to encourage point estimation practice. The game is full screen, with a number range of 0 to 9. Players move the mouse to listen for ships and numbers, and press the number that is near the ships. In this example, pressing “7” would destroy 3 ships and give six points, while pressing “5” would not hit any ships and cost 1 point.

ships.

Ships were released using the following rules:

1. Each ship is released randomly from the left or the right side of the screen.
2. There is a maximum number of ships on screen.
3. The release rate for ships begins at 0.2 per second (1 every 5 seconds).
4. The release rate increases by 0.05 per second after each minute-level.
5. The release rate “burst” to five times the baseline rate during the first 20 seconds.
6. The velocity of the ships varies (some are faster than others).

Scoring was 2 points for each ship hit, -1 point for each miss, and -1 point for each ship that crosses the entire screen. When a player selected a marker with more than one ship, she blew up all of the ships at that marker. There was a maximum of one shot per second, with no penalty for firing early (but it would not function). Therefore, the number of minimum points for each level was -60, less the number of ships for that level. The maximum number of points for each level was twice the number of ships in the level. Based on the release rules above, the first wave would have 48 ships, and each additional wave would have 12 more ships. Table 11 shows some potential scores for the Navy game.

Table 11: Scores for the Navy game.

| Level | Ship Count | Min. Score | Max. Score | Cumulative Max. Score |
|-------|------------|-------------------|----------------|-----------------------|
| 1 | 48 | $-60 - 48 = -108$ | $48 * 2 = 96$ | 96 |
| 2 | 60 | $-60 - 60 = -120$ | $60 * 2 = 120$ | 216 |
| 3 | 72 | $-60 - 72 = -132$ | $72 * 2 = 144$ | 360 |
| 4 | 84 | $-60 - 84 = -144$ | $84 * 2 = 168$ | 528 |
| 5 | 96 | $-60 - 96 = -156$ | $96 * 2 = 192$ | 720 |

Like AudioFitts, the Navy game was also designed for logging user actions. Each display frame reported the time, the position of the ships, and the users score. As shown in Phase 3 (Chapter 6), students showed improvement in the Navy game over time, suggesting that they were also improving at APE.

With the basics of APE in place with AudioFitts, and training ready with Navy, the previously unknown factors of an accessible graphing technology for classes were uncovered, and development on the final system would begin.

3.6 GNIE: Active Point Estimation In Classrooms

The final technological challenge was to make an accessible graphing tool for low vision and blind students. The AGE provides a way to store graphs and convert them to various formats. The AudioFitts studies completed in Phase 1 demonstrated how to make APE possible with audio. Together they inform the design of the final tool, called Graph and Number line Interaction and Exploration system (GNIE).

The GNIE user interface has three main parts. Most of the application is a Java program with typical interaction widgets, specifically menu items, check boxes, text boxes, and buttons. These widgets inherit accessibility hooks automatically, and were enhanced with additional accessibility context for screen readers. The second part is a text area, covering the left half of the screen. This holds any non-graph parts of a graphing question¹³, such as the text "Graph $(-2, 3)$." The third part is a graph (or number line), covering the right half of the screen¹⁴. This space was designed to follow the APE settings found in the Phase 1

¹³The AGE was expanded to hold this text, so that the graph question and/or answer could be stored with the graph.

¹⁴To manage mouse movement a little more easily, and ensure sufficient graph movement space, the program automatically maximizes its size. The right half of the application, then, is also the right half of the screen.

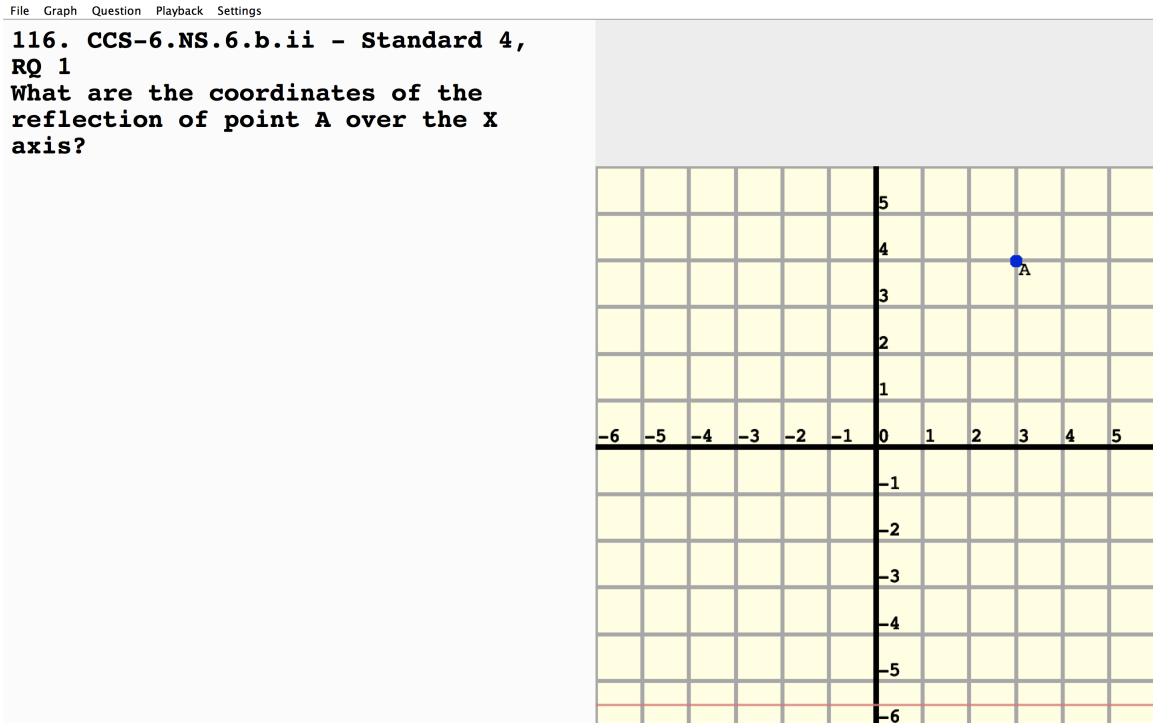


Figure 28: The GNIE program.

AudioFitts studies; all accessibility was handled explicitly by GNIE, to control for screen reader effects. Figure 28 shows an overview of the program.

Activity logging is more limited in GNIE. Since the software is meant to be used in classrooms, evaluation focused on live classroom use (screen captures and third person video) and the accuracy of the user's final work.

3.6.1 Future Work

GNIE was sufficient to complete the tasks for this dissertation (discussed in Phase 3). However, longer term evaluations in classrooms are still necessary. Work in the Sonification Lab in this area continues, and there is a positive outlook on further improvements.

3.7 General Discussion

When designing graphing tools, it is tempting to think about graphing in the only visual sense, but that approach is problematic. The goal of graphs is to communicate data in a form that highlights the relationships of the two variables. It is perfectly acceptable to have part of that interaction occur in a non-visual mode.

Those readers in the accessibility fields may be wondering: why are haptics missing from this conversation? I made a deliberate effort to *avoid* the accessibility advantages of haptic and tactile displays for many parts of these programs. The reason is purely scientific: findings that point to users being able to graph will be due to auditory feedback, and not the haptic or tactile feedback¹⁵. Based on the findings in Phases 1, 2, and 3, it appears that audio APE is useful. One future direction should be to combine the audio work with touch, using absolute positioning devices such as touchscreens, or other tactile assistive technologies.

Table 10 shows 19 variables that were set in each AudioFitts experiment. Many combinations from this list have yet to be explored, and other important variables have not been evaluated at all. There is a long line of studies that are necessary for fully understanding audio APE. The accuracy of point estimation results when using this method is established with this dissertation. Audio, however, is about three times slower than visuals. Research targeting the cause and possible solutions to this problem would benefit audio APE and accessible graphs research¹⁶.

A final important improvement would be on the auditory design. GNIE and the other programs have proven to be usable by students and teachers, but the selection of sounds has drawn its share of chuckles and grimaces. Fine tuning the audio design will lead to a higher user comfort and probably higher performance.

3.8 Conclusion

Accessible graphs had many tools before this dissertation. They simply have not been part of a cohesive development effort. The AGE brings a model (the AGM) and a format converter (the AGFC) to accessible graphs. AudioFitts presents a way to test APE, and Navy lets students learn how to apply APE. Finally, GNIE delivers graphing possibilities to students, enabling them to use APE to solve their everyday graphing problems.

¹⁵The mouse and keyboard certainly have tactile and proprioceptive feedback, but are not designed to give any accessibility advantages. Important keyboard keys were marked with locator dots.

¹⁶Based on my discussion with teachers and test administrators, adding time for visually impaired students is not difficult; in many cases it is standard practice. Therefore, a longer time factor of even 3 may be acceptable. I suspect that through user training and stimulus improvements, students can learn to be much quicker than they were in the tests reported in this thesis.

CHAPTER IV

PHASE 1: POINT ESTIMATION WITH AN AUDITORY DISPLAY

Blind and low vision students may be able to use auditory graphs for mathematics education. The first step is to establish a way to use sonification for point estimation (phase 1). In phase 2, a method for identifying graph-based system requirements is presented. In phase 3, I propose a series of studies for system evaluation in classroom and testing situations.

Point estimation is a critical component to graphing, required in almost every graphing question. However, it is not clear how a user could use non-speech audio to estimate a point, especially using the standard keyboard and mouse. The interaction device used and the way the sound is displayed required further basic research before application into a system. These studies were completed between August-December 2011, are summarized in table 12, and in more detail below.

4.1 Study 1: The effect of input device on target speed for visually impaired participants.

On a computer, sighted students generally use a mouse to navigate a graph. However, many visually impaired students do not have the necessary feedback to use the mouse, so they are limited to using the keyboard. This first study established whether visually impaired people can use sonification and either the keyboard or the mouse for a targeting task related to point estimation. This study presented a one-dimensional targeting task, similar to finding a point on a number line. This targeting experiment is related to Fitts' Law experiments

Table 12: The four studies in phase 1. These establish how to use sonification for point estimation.

| Study | Description |
|-------|--|
| 1 | The effect of input device on target speed for visually impaired participants. |
| 2 | The effect of input device and presentation format on target speed for sighted participants. |
| 3 | The effect of auditory scaling and peaking on target speed and accuracy. |
| 4 | The effect of auditory scaling and mapping type on target speed and accuracy. |

Table 13: The effect on the cursor position from pressing the 10-key number pad. The 'none' column of numbers have no effect on movement.

| Step Size | Direction | | |
|--------------|-----------|------|-------|
| | Left | None | Right |
| 100px | 7 | 8 | 9 |
| 10px | 4 | 5 | 6 |
| 1px | 1 | 2 | 3 |

(e.g. [? ? ?], discussed in more detail in in Section ??).

There were a few differences from study 1 and other Fitts's Law experiments. First, the population in study 1 was visually impaired instead of sighted. Study 1 also presented a sonification instead of a visualization. In terms of independent variables, the keyboard input and the target distance were modified as described below.

Card et al. [?] modeled the movement time for keyboard after the minimal number of key presses necessary to reach the target. However, it would be useful to somehow model the keyboard in terms of target distance and target size, so that the keyboard and mouse can be compared. The model of key movements used is similar to the "Text Keys" described in [?], except the user is moving by pixel instead of letter steps. In this study, pressing a particular key moved the cursor 1, 10, or 100 pixels to the right or to the left. The spatially lowest keys on the 10-key number pad moved the cursor 1 pixel, the middle keys moved 10 pixels, and the highest keys moved 100 pixels. The left keys moved the cursor left and the right keys moved right. So, for example, the '4' key moved the cursor 10 pixels left, and the '9' key moved 100 pixels to the right. See Figure 29 for a picture of the 10-key number pad and Table 13 for an overview of the effect of each key.

In pilot studies, there was a wide variance of movement times for similar target distances. This only occurred when using the keyboard. Similar to Card et al. [?], the keyboard time is highly dependent on the number of keys pressed. A target width of 5 pixels and target distance of 680 pixels from the target takes more time than the slightly longer target distance of 700 pixels because it takes more key presses to reach 680: $7 * 100px - 2 * 10px = 680$ pixels and 9 keys¹, versus $7 * 100px = 700$ pixels and 7 keys.

¹In this case, it takes fewer keys to move to 680 by first moving to 700 then moving back 20 pixels to 680. If the model only allows moving toward the target, it would take $6 * 100px + 8 * 10px = 680$ pixels and 14

In pilot studies, it also became clear that switching between the keys for different size jumps (from 100px to 10px to 1px) produced large effects on movement time. Beyond simply counting the number of keys, perhaps counting the key switches necessary could be used to predict movement time. In addition, the definition of “key count” is unclear. It could be the fewest number of key presses, the number of presses without passing over the target, the actual key count, or some other measure. Ideally, the time will be predictable, and in that case the model must be predictable, so the actual key count is not an option. This study used two strategies to make the keyboard results comparable to the mouse results. First, it introduced a new measure, key level count, to explore whether required key switching is a prediction of movement time. Key level count could be controlled for a more even comparison with mouse results. In addition, instead of a fixed distance, the study used a bounded range of distances, so that the shorter and longer number of key presses would average into the typical time for keyboard movement.

The new measure, key level count, is based on the depth of the level that a user would have to go to in order to reach a target, starting at the highest jump size. Participants want to be accurate and fast. They will likely start with the largest jump. Based on pilot studies, participants will move to the medium and small jumps only when it is necessary to reach the target.

The peaks and valleys due to key count and key level count can be smoothed by using a range of target distances between targets, instead of a fixed target distance. The eventual point estimation tasks on graphs and number lines require students to move arbitrary distances to make points². It is reasonable to evaluate the effect of distance, but this cannot be done for fixed distances due to the non-continuous nature of key presses over small changes in distance. Evaluating the average movement time for a range of target distances will result in a comparable measure between the mouse and keyboard. For the index of difficulty calculation used in Fitts’s Law studies, $ID = \log_2 \left(1 + \frac{D}{W}\right)$, the target

keys. In pilot studies and the experiment, participants often went over the target and returned, with both the keyboard and the mouse.

²Tick marks may be evenly spaced, but the space between tick marks is based on the available screen space and the size of the graph, not a fixed pixel width. In addition, people might plot points between tick marks. Therefore, the tool should enable easy navigation to arbitrary distances.

distance could be averaged to a point. In order to have the same key presses between ranges, other than the key presses it takes to get to the range, ranges should be a factor of the same size as the largest key jump. In this case, 100 pixels is the largest jump size, so 300 pixels was used as the range, starting at 20px. The ranges can be shifted, but should also ideally begin at the same shift. Since target distance ranges have different averages, they can be compared in terms of index of difficulty and effect on movement time. A larger average target distance should take more time to get to simply because it takes key presses to get to the range (e.g. more key presses to get to the 620-919 range than the 20-319 range).

There were the following questions for Study 1:

Q_1 What is the effect of target distance range on keyboard movement time?

Q_2 What is the effect of target size on keyboard movement time?

Q_3 What is the effect of target distance range on mouse movement time?

Q_4 What is the effect of target size on mouse movement time?

Q_5 What is the effect of input device on movement time?

Q_6 What is the effect of key level count on movement time?

Q_7 What is the effect of level of vision impairment on movement time?

The Fitts's Law questions constitute a single family of statistical tests, and alpha should be adjusted accordingly. The key level count and vision impairment tests are independent and were be tested on $\alpha = 0.05$.

There were the following predictions for the keyboard. A larger target distance range results in a larger movement time. A smaller target width results in a larger movement time. In addition, a higher key level count results in a higher movement time. Finally, there may be a positive linear relationship between the index of difficulty and movement time for the keyboard condition.

Mouse movement time should increase when target width decreases or target distance range increases. There should also be positive linear relationship between the index of

difficulty and movement time for the mouse. Since the eventual application is for number line and graphing tasks, target widths are small and distances can be large. Thus the index of difficulty is moderate to high (e.g. above 3). In addition, the study will use non-speech audio feedback, which Edwards [?] has shown to be somewhat slow. The linear model of Fitts's Law therefore appears to be appropriate³. It was not known what to expect for the impact of input device, but tested for a difference. In visual interfaces, the mouse has been faster than the keyboard [?]. In accessibility applications, the mouse is often not used, and has had slow movement times [? ?].

In addition, the experiment evaluated whether self-reported low vision participants are faster or slower than blind participants. A reasonable expectation is that blind participants will have higher movement times (they will be slower) than low vision participants, since they don't normally use the mouse and since they don't normally get continuous feedback for movement. The study also collected functional impairment and other demographic relationships to the level of visual impairment. Since the eventual application is targeting visually impaired users, visually impaired participants were the subjects for the study (as Edwards suggests [?]).

4.1.1 Study Design

This study balanced considerations of lab and ecological evaluations. While the environment was controlled, the study included realistic design decisions rather than ones that may control the situation more. For example, mouse acceleration remained at its default for the computers being used. A Fitts's Law study, for example, could remove software acceleration because it impacts the timing results. However, it would be less realistic to find a mouse with disabled acceleration. That said, the specific mouse used and the specific mouse acceleration was the same for every computer.

³For very quick reactions, Fitts's Law may not be appropriate [?]. Note that Fitts's Law is logarithmic with respect to target distance and target size, and linear with respect to index of difficulty.

Figure 29: The mouse and the right side of the keyboard used in the experiment. Note the additional keys around the number keys, and the horizontal tactile mark on the 5 key.



4.1.1.1 Apparatus

The experiment space was at a small computer lab at a local visual impairment work rehabilitation organization, the Center for the Visually Impaired, in Atlanta, Georgia. Participants were seated about 4 feet next to each other. Participants could adjust their chair and general workspace before the study and during training.

The computers were dual core CPUs, with a range of 2.13 to 3.0 GHz, and 2GB RAM. The operating system was Windows 7 Professional, 64-bit. The study used Dell SK-8175 keyboards and Dell MS111-L mice (both use USB cords). Participants wore Sennheizer HD 202 headphones and sleep shades.

Participants used the spacebar to select, and the 10-key number pad, specifically numbers 1,3,4,6,7, and 9 to move (See table 13 for a description of movements). The numbers worked on the system regardless of num-lock status. The keyboard in the experiment, like many other keyboards, has several keys surrounding the number pad, and has a small tactile mark on the “5” key for non-visual orientation. See Figure 29.

There were specific settings for the keyboard and the mouse set in the Windows Control Panel. For the keyboard, repeat delay was set to 3 (of 4), and repeat rate was set to 32 (of 32). There were several more changes for the mouse settings. Mouse motion was set to 6 (of 11). “Enhanced pointer precision” was on. Snap was off. Trails were off. Hide pointer while typing was on. Showing location of pointer when control key is pressed was off. Double-click speed was 7/11. Click lock was off. Mouse acceleration was on. The mouse and the keyboard settings were based on the most common settings for the computers in the lab.

The AudioFitts program was built with Java, with the processing and Minim libraries⁴. Part of the motivation for using Java is based on compatibility with current programs for accessible graphs and number lines that use Java (such as the Sonification Sandbox). Processing and Minim also made it easier to build a lightweight system. However, there were notable audio challenges, in part because of the choice in language. There was choppy audio and high latency on certain machines. In order to address these concerns, I adapted the way the system displayed the audio until the latency issues decreased to the point that it was difficult to detect. Nevertheless, one participant commented on the audio latency. Based on latency reporting from the Java Sound engine, the estimated latency was about 70 milliseconds. However, I am not confident that this latency is completely accurate, as that is the reported guaranteed upper bound of Java’s latency. Future work should gauge latency better. The functional impact of latency on audio-based tasks appears to be an open area of research.

There were 3 key levels, of 1, 10, and 100 pixels. A quick analysis of potential trios of numbers to use showed these were close to optimal values for the full range of potential distances between 20 to 920 pixels. In addition an equal factor gap (of 10) between the levels may provide to a simpler mental model for participants. In addition, all steps divide into the target distance range of 300, so the average key count for each target distance range is the same, if the key presses to the range are ignored.

⁴The processing library can be found at <http://processing.org>. Minim is at <http://code.compartmental.net/tools/minim/>.

4.1.1.2 Stimuli

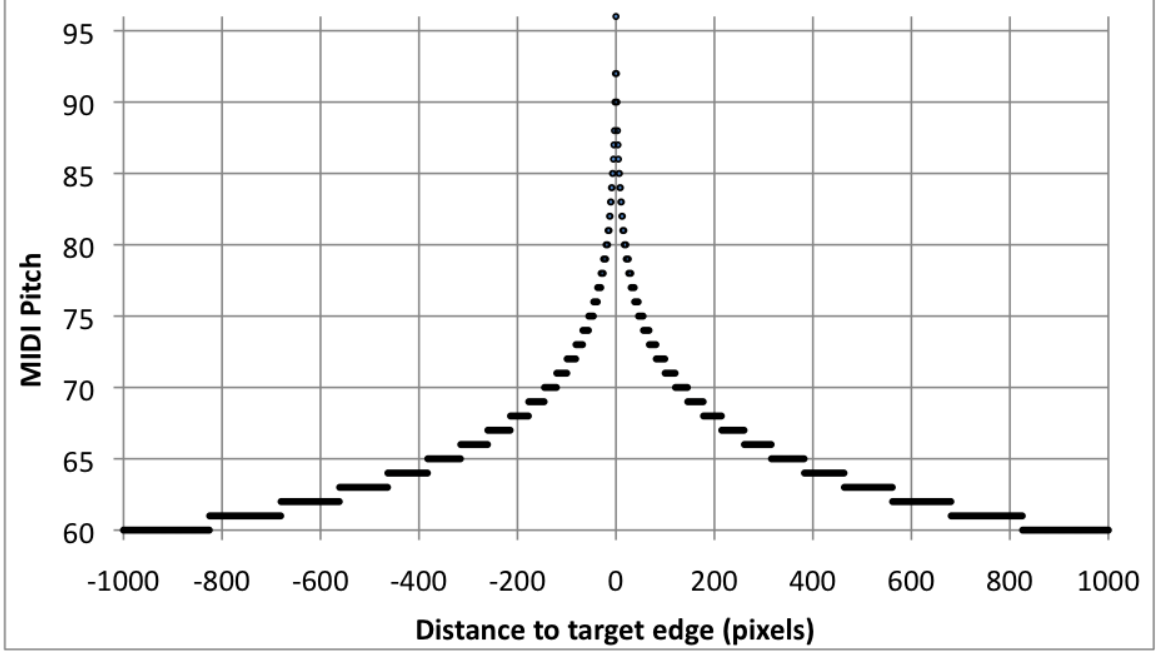
The mapping of left-right direction was mapped onto pan, timbre, and rate. There were slightly different timbres (in MIDI, different musical “organs”) for being left or right of the target. If the target was to the left (the cursor was to the right of the target), the sound came out of the left ear; if the target was to the right, the sound came out of the right ear. An alert sound indicated when the user had the cursor on the target. The system also indicated when the user “jumped over” the target (e.g. the cursor skipped over the target from the left of the target to the right of the target). Thus, there are four sounds for direction: left target, right target, on target, and passed over target. The on target stimulus sounded like quickly switching back and forth between two piano keys, an octave apart (MIDI notes 84 and 96). The passed over target stimulus was the same rate as the on target stimulus, an octave above the other keys (MIDI notes 96 and 104). The passed over target sound was played for a single applet frame, typically about 50ms, while the on target sound continued while the participant remained on the target. This design provided similar themes but noticeable differences. The on target sound and passed over target sound are effectively alert earcons.

In addition to direction, the mapping of current distance to target was mapped onto pitch. The pitch value of the distance sound was related to the distance. A closer distance resulted in a higher pitch. Initially, sine tones were used, mapping distance to pitches within the frequency range of 200-2000Hz⁵, in the middle of human hearing range [?]. The display later shifted to MIDI-based tones (with no pitch bending), within the same approximate frequency. While there is a disagreement whether sine-based or MIDI-based displays should be used (e.g. [? ?]), people can distinguish pitch better with MIDI [?].

In pilot studies, it felt as if there needed to be more information when close to the target, and the pitch changes with small mouse and keyboard movements (for sine and MIDI) wasn’t enough. To fix this, the value-to-pitch mapping was changed from from

⁵These frequencies were adjusted to a log scale so they map to the human perception of frequency, known as pitch [?].

Figure 30: The relationship between the distance to the target edge and the MIDI pitch, using the logarithmic scale. When inside the target, the user hears an alert. See Equation 5 and Table 17 for examples on the effect of the logarithmic mapping.



linear to logarithmic. Given absolute pixel distance D , screen width S ⁶, lower MIDI pitch value L , and higher MIDI pitch value H , the final distance-to-MIDI note formula used was:

$$N = \lfloor H - (H - L) \left(\log_{10} \left(999 * \left(\frac{D}{S} \right) + 1 \right) \right) / 3 \rfloor \quad (5)$$

N is the value of the final MIDI note, scaled to the note range specified for the program and concatenated into an integer. For this study, L is MIDI note 60 (C4, 261.6 Hz), H is MIDI note 96 (C7, 2093.0 Hz). Within the first few pixels, many MIDI notes are used. Further out, a single MIDI note is used for a few distances, and even more as the distance increases. This produces a balance of many note changes near the target and fewer note changes when further from the target. In other words, this mapping gives more detail near the target at the expense of using up the 37 available notes. Multiplying the ratio by 999 and adding 1 produces a range of 1 to 1000⁷. The 999 multiplier is removed after

⁶The equation looks similar to Fitts's Law, but it is intended for a different purpose. S in this case is screen width, not target width.

⁷less than 1 in results negative log values, to negative infinity)

the log transformation with a division by 3 seen in the second formula ($\log_{10}(999)/3 \approx 1$). The resulting log scale, shown in figure 30, produces a lot of detail near the target, while maintaining a little auditory feedback at locations further away from the target.

There was a lot of trial and error in designing the auditory display. There was insufficient empirical data for designing auditory displays that support point estimation. In several ways, the display was a “best guess”. The feasibility of the mouse and the keyboard as potential devices for visually impaired users had to be established first, so that subsequent designs could focus on one or both interaction devices. The empirical research to establish the auditory display design was conducted in studies 3 and 4. As it turned, a pitch-based logarithmic mapping was one of the best candidates.

4.1.1.3 Participants

Participants were recruited with the help of staff at the Center for the Visually Impaired (CVI), and consisted of organization staff, clients, and local community members who self-identified as visually impaired.

Of the 25 participants who were included in the data analysis, 17 were male (8 female), 22 were right handed (3 left handed), and all 25 reported normal hearing. The average age was 44.04, with a standard deviation of 11.781, minimum of 21, and maximum of 61.

Participants gave self-reported and functional levels of vision impairment, sometimes conflicting on initial analysis. Self-reported level of visual impairment was 10 blind, 15 low vision. 5 reported being visually impaired for less than 10 years, 10 for more than 10 years but not their entire life, and 10 reported being visually impaired for their entire life. We did not ask how long participants were low vision or blind, only if they had visual impairments. In terms of functional visual impairment for desktop computing, 12 used screen readers, 2 used magnifiers, 8 used both, and 3 used neither assistive technology. 10 reported currently using the mouse when computing (15 do not currently). Self-reported blindness did not often fully match with expected functional effects of blindness. For example, of the 10 who reported being blind, two reported using the mouse (8 reported not using the mouse) and one reported using a screen reader and magnification software (9 reported not using

magnification software). Their reported descriptions of mouse use explains why: one uses it for high-level navigation, The other uses it for his touchscreen phone. Both were referring to Apple products using the VoiceOver technology as a screen reader.

The participant sample generally had experience with computers and the keyboard, and 64% with the mouse. Only 2 participants reported having less than 3 years computer experience; the rest (23) had 5 or more years. Similarly, only 3 participants reported having less than 3 years keyboarding experience. 12 had used a computer before their visual impairment, 12 had not. 15 do not currently use a mouse, 10 use a mouse. 16 have one or more years of mouse experience, 9 have no experience (all of those with no experience were blind). 5 had used a mouse before their visual impairment.

4.1.1.4 Task

For each experiment session, 1 to 5 participants were given information about the study and signed informed consent with a reader and witness. Participants were then led to the computer lab and seated at a computer. Everyone was asked to put on sleep shades, followed by headphones. The experiment team then described how the keys and mouse would be used for the experiment, and gave time for participants to adjust their seating and input devices.

The study began with a spoken paragraph of overview instructions, from a prerecorded audio clip. The instructions explained what the target and distance sounds were like. The participant then completed two sets of blocks, run as keyboard then mouse or mouse then keyboard (counterbalanced between participants). At the start of each session, participants completed a training, which described how to use the interaction device to complete the task along with blocks with increasing difficulty (smaller target widths and higher distances). Note that some participants may have never used a mouse so this basic training could be crucial. For the first five blocks, participants were required to hit a fixed number of targets (regardless of time), with non-speech auditory feedback indicating a hit or a miss. The last two blocks simulated the evaluation sessions, each lasting 45 seconds (regardless of hit count) and without positive or negative feedback. Research staff were encouraged to assist

participants who asked for help or appeared to be struggling. Staff did not assist with targeting during the experiment portion of the session.

After training, participants completed 9 blocks. Each block was 5, 9 or 13 pixels wide, with a target distance range of 20-319, 320-619, or 620-919 pixels. The order of the 3x3 blocks was randomized by the AudioFitts system for each participant. After the first set of blocks, the experiment team started the second session, starting with training for the particular interaction device.

Between blocks, participants were asked by a prerecorded voice to select the target to begin. Once participants selected that target, the trials started (and the voice stopped talking). Each trial lasted as long as it took for the participant to complete the task (with a maximum of the block's time length). Like the original Fitts' study [17], we chose to have a fixed time for each block, instead of a fixed number of trials per block. This allows each participant the same amount of time for the session, which was important for the time constraints at the computer lab. In addition, participants would each have the same amount of fatigue in terms of time on the task.

During a block, participants heard the distance to the target. When they were on the target, they heard the on-target sound. They attempted to select the target. Regardless of a hit or miss, the participants heard the next target.

Blocks lasted 45 seconds with a user-selected start for each block. This makes the study a "continuous" Fitts' Law task (the targeting is continuous from trial-to-trial. Note, however, that target distances changed in a range for each trial. Participants could not mechanically "memorize" the distance due to these variations. Therefore, mouse and keyboard movement times are expected to be higher than in other Fitts's Law studies that use fixed, single distances. However, the varying distance is more realistic to the plotting task and, as described above, allows for keyboard comparison with target distance (range) and target width. The experiment part of each block set (9 blocks) lasted between 7-10 minutes per participant, with the training lasting 5-10 minutes per participant. For two block sets (mouse and keyboard), the time on the computer ranged between 30-40 minutes.

After completing the experiment sessions, participants were led to a larger room. While

Table 14: Study 1 statistical tests conducted. Effect size is for μ_p^2 (partial eta squared) for the Repeated Measures ANOVA, and Cohen’s d for Student’s t-test. Contrasts are linear. Horizontal lines indicate families of tests.

| Topic | Device | Result | p-value | Effect Size |
|----------------------------------|----------|---------------------|--------------|-------------------|
| Target Width | Keyboard | $F(2, 36) = 16.568$ | $p < 0.0055$ | $\mu_p^2 = 0.479$ |
| Target Width | Mouse | $F(2, 48) = 23.862$ | $p < 0.0055$ | $\mu_p^2 = 0.499$ |
| Target Width, Contrasts | Keyboard | $F(1, 18) = 49.483$ | $p < 0.0055$ | $\mu_p^2 = 0.733$ |
| Target Width, Contrasts | Mouse | $F(1, 24) = 60.789$ | $p < 0.0055$ | $\mu_p^2 = 0.717$ |
| Target Distance Range | Keyboard | $F(2, 36) = 15.984$ | $p < 0.0055$ | $\mu_p^2 = 0.470$ |
| Target Distance Range | Mouse | $F(2, 48) = 20.573$ | $p < 0.0055$ | $\mu_p^2 = 0.462$ |
| Target Distance Range, Contrasts | Keyboard | $F(1, 18) = 24.156$ | $p < 0.0055$ | $\mu_p^2 = 0.573$ |
| Target Distance Range, Contrasts | Mouse | $F(1, 24) = 32.371$ | $p < 0.0055$ | $\mu_p^2 = 0.574$ |
| Input Device | Both | $F(1, 24) = 19.681$ | $p < 0.0055$ | $\mu_p^2 = 0.451$ |
| Key Level Count | Keyboard | $F(2, 42) = 6.768$ | $p < 0.025$ | $\mu_p^2 = 0.244$ |
| Key Level Count, Contrasts | Keyboard | $F(1, 21) = 8.262$ | $p < 0.025$ | $\mu_p^2 = 0.282$ |
| Level of Vision Impairment | Keyboard | $t(23) = 1.316$ | $p = 0.201$ | - |
| Level of Vision Impairment | Mouse | $t(23) = 2.846$ | $p < 0.01$ | $d = 1.04$ |

eating pizza and collecting a \$10 check, participants filled out a demographics survey. The overall experiment lasted about 50 minutes per participant.

4.1.2 Results

Of the 35 participants that were in the study, 25 completed the sessions. 5 were excluded because they could not complete the study. Non-completion did not appear to have an association with a particular device or level of vision impairment. In addition, 5 participants had not finished a pilot that was aborted due to technical problems. We allowed 4 to participate in the full experiment since this study is working with a population that is difficult to recruit and because the analysis was largely done within subjects. One participant (included) had paralysis in his left arm.

Evaluation of the interaction devices and impact on vision level were the only open questions we had about the direction of the differences. Target width and target distance (range) effects on time are fundamental to five decades of Fitts’ Law research; we expected this to hold for auditory displays (as it has in similar works [7]) for the mouse. Key presses and time is fundamental to KLM and other GOMS [14] models, and so we expected a relationship with time and keyboard press difficulty.

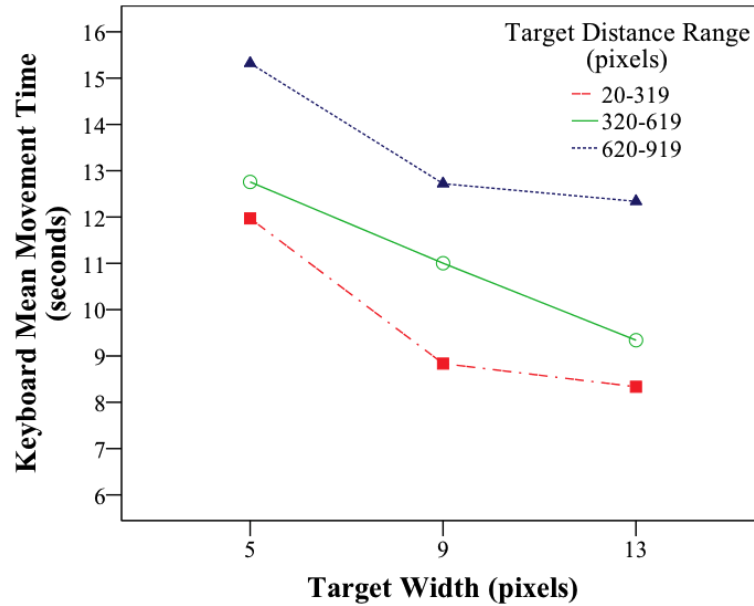


Figure 31: Target width, target distance range, and keyboard mean movement time.

4.1.2.1 Keyboard

The mean movement time for each keyboard trial was 10.030 seconds (SD=6.176). The keyboard hitrates had a mean of .9348 (SD=.09264; minimum=.64).

All tests with the keyboard showed significant differences. All tests were conducted with a within subjects ANOVA, with contrasts. See table 14 for statistics and figure 31 for a comparison of target distance, target width, and time for the keyboard.

Movement times for target width are significantly different. The shortest times were on average for the larger target widths. The mean difference between the smallest target width, 5, and the largest target width, 13, was 3.201 seconds (SD=4.956).

Movement times for target distance range are significantly different. The shortest times were on average for the smaller target widths. The mean difference between 620-919 pixels and 20-319 pixels was 4.668 seconds (SD=5.383).

Movement times for keyboard level count (KLC) are significantly different. The shortest times were on average for the fewer keyboard levels. The mean difference between the KLC of 3 and 1 was 4.031 seconds (SD=6.577).

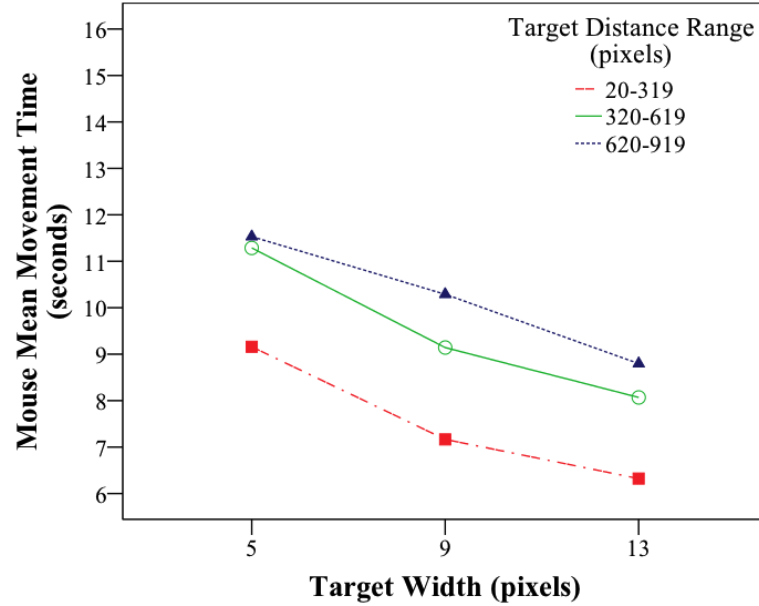


Figure 32: Target width, target distance range, and mouse mean movement time.

4.1.2.2 Mouse

The mean movement time for each mouse trial was 7.560 seconds (SD=4.005). The mouse hitrates had a mean of .9260 (SD=.04318; minimum=.83).

All tests with the mouse showed significant differences. All tests were conducted with a within subjects ANOVA, with contrasts. Key level count was not evaluated for the mouse. See table 14 for statistics and figure 32 for a comparison of target distance, target width, and time for the mouse.

Times for target width are significantly different. The fastest times were on average for the larger target widths. The mean difference between the smallest target width and the largest target width was 2.568 seconds (SD=1.666).

Times for the target distance range are significantly different. The average times were faster for shorter target distance ranges. The mean difference between 630-919 pixels and 20-319 pixels was 2.685 seconds (SD=1.926).

4.1.2.3 Input Device and Vision Impairment

Participants were significantly faster with the mouse over the keyboard. There are two ways to compare the statistics, either per trial or per block. Per trial, the mean difference in

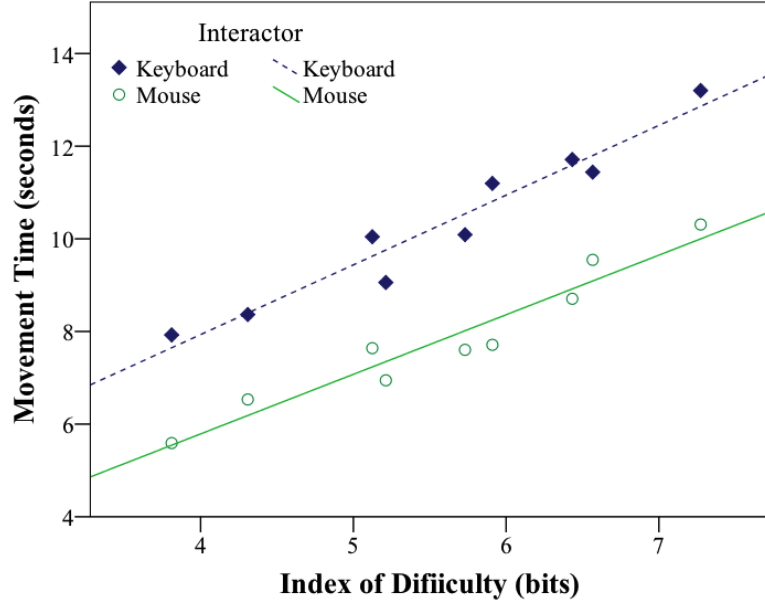


Figure 33: Index of difficulty and movement time for the mouse and the keyboard.

trial movement times was 2.470 seconds. Keyboard mean was 10.030 seconds, median 8.406 seconds (SD=6.176, n=779 trials). Mouse mean was 7.560 seconds, median 6.675 seconds (SD=4.005, n=1111).

Faster participants completed more trials per 45-second block, so are over-represented in a per-trials comparison⁸. Per block, the difference is greater. Keyboard mean time was 12.728 seconds (SD=6.957 seconds, median 10.945 seconds). Mouse mean time was 8.544 seconds (SD=2.842 seconds, median 8.088 seconds). The keyboard was therefore 49% slower than the mouse based on the mean, and 35% slower based on the median. It will take 4.2 seconds less time on average for a visually impaired person to find a target with the auditory display using a mouse instead of a keyboard.

Participants who self-reported as blind (n=10) were significantly faster than low vision participants (n=15) for the mouse ($t(23) = 2.846$, $p=0.009$; mean difference=2.430 seconds, standard error of the difference = 0.854, $d=1.192$). There was not a significant difference between the keyboard for low vision and blind participants ($t(23) = -1.316$, $p=.201$). Although their descriptive means differed by 3.055 seconds (blind faster), the standard

⁸All statistical tests used a per-block comparison so that participants would be compared equally.

error of the difference was 2.321 seconds. In general, keyboard times varied much more than mouse times.

Self-reported blindness was not significantly correlated with gender, handedness, or age⁹. Therefore, these factors do not appear to be an underlying cause of the blind participants' higher performance.

Figure 33 shows the index of difficulty plotted with the movement time for the mouse and the keyboard. For the keyboard, the fit line equation is $MT = 1.5054 * ID + 1.9136$, $R^2 = 0.942$. For the mouse the fit line equation is $MT = 1.2869 * ID + 0.642.15$, $R^2 = 0.928$. The fits are reasonable, especially considering the use of target distance range, auditory displays, and visually impaired user groups (all different from traditional Fitts's Law studies). It is remarkable that the keyboard's coefficient of determination is higher than the mouse, considering that the keyboard has typically been excluded from Fitts's Law modeling of target width and target distance effects on speed.

4.1.3 Discussion

Based on these results, there are the following conclusions. For a Fitts's Law targeting study using sonification with a logarithmic distance mapping:

1. As target distance range increases, movement time increases for the keyboard and the mouse (Q_1, Q_3).
2. As target width increases, movement time decreases for the keyboard and the mouse (Q_2, Q_4).
3. The movement time decreases when using the mouse instead of the keyboard (Q_5).
4. The movement time decreases when using fewer key levels (Q_6).
5. Blind participants are faster than low vision participants with the mouse (Q_7).

Early work with mice in non-speech audio suggested that interaction was relatively slow, on the order of 4-7 seconds [9]. In the current study with AudioFitts, the task of finding

⁹These tests were accounted for in a post-hoc Bonferroni Correction of the familywise alpha from 0.025 (2 tests) to 0.01 (5 tests).

a small target over a large target distance range had index of difficulty values ranging 3 to 8. Based on results from other studies ([1,16]), the mouse slope is 3 to 13 times slower when using an auditory display. However, these studies had somewhat different designs than the current work. Replicating and extending this study to include sighted participants in visual, auditory, and combined conditions using an enhanced AudioFitts would provide the best comparison metric. This was completed in study 2.

Visual displays are sometimes faster than auditory displays. However, the mouse may be the fastest device available for low vision and blind computer users doing certain tasks if only a regular desktop computer is available. In terms of testing accommodations, the mouse and the keyboard are likely to be acceptable tools for sighted students (if a computer is acceptable), so it may be easier to have them approved for visually impaired participants.

With the target distance range, the keyboard fit remarkably well into the Fitts's Law model. It may be useful to use the target distance range approach when comparing key devices with continuous devices. Further work on predicted key count, actual key count, key level count, and using fixed target distances and target distance ranges would clarify the best predictors of keyboard movement time.

4.1.3.1 Measures for Key-based Inputs

Key-based inputs such as the keyboard have been treated as fundamentally different from continuous input devices such as the mouse, as early as Card et al. [?]. Continuous input devices have a Fitts's Law model, while key-based devices have a key press model. In order to verify these models of human movement, more work should be done to compare the models with the data available on mouse and keyboard movements. The keyboard model is particularly under-explored. While KLM [?] (and similar) models can capture intentionally discrete actions, such as selecting a menu item, they may not be as suitable for modeling human movement with canvas-like spaces such as number lines and graphs.

One major challenge is identifying how to count predicted key presses. One obvious approach is to count the number of keys it takes to get to, and not over, the target. Another way is to jump over the target only if it takes fewer key presses. In pilot studies,

the experiment team found ourselves jumping over the target multiple times, then deciding to change key levels. Our proposed key level count captures the level switching, but not the key presses to reach a point. There may be other straightforward models as well. Naturally, these models should be compared with both movement time and the actual key presses, delay time, and when participants switched key levels. This work would extend GOMS-KLM to visually impaired users (like [?]) and for canvas-like spaces.

4.1.3.2 Visual Impairment and Interaction Devices

Blindness is a condition that is physiological, functional, and social. Of the 25 participants included in analysis, 10 reported blind, 20 use screen readers and 10 use magnifiers. Also note that 10 of 25 use the mouse. The level of mouse use is a useful indicator of functional level of visual impairment considering the purpose of this study. However, these indicators did not completely match the self-reported description of blindness. “Blind” is specific to the task and program of research. In addition, there is not sufficient agreement on how to measure blindness, functional or otherwise. It is therefore difficult to compare results with other studies. Future work could explore the relationships of a variety of survey questions on blindness.

It appears that the visually impaired, whether low vision or blind, can use the keyboard or the mouse for 1-dimensional tasks, such as putting a point on a number line.

It is surprising that blind participants were faster than low vision participants. Perhaps blind computer users are familiar with auditory displays and may be more skilled at interacting with them, regardless of input device. It would then follow that there would be larger differences in the keyboard, since blind people often have more experience with the keyboard. The keyboard, however, did not have significant differences (there was a 3 second descriptive difference, faster for the mouse). Further work with the mouse and other interaction devices with visually impaired users will show more clearly the differences, and may lead to a stronger theory.

About 5 people had trouble initially orienting themselves with the number pad. Problems included knowing which keys were number keys, and that the number pad is reversed

from a typical phone number pad (1 is on lower left on the number pad, instead of upper left like the telephone). Figure 29 shows the keyboard; keys to the top, right, and bottom of the direction keys confused a handful of participants, particularly during the training. The 5 key on keyboard has a tactile mark; however some participants had trouble feeling this mark. Classroom research later showed that tactile locator dots may be more suitable.

Participants used the right hand to move and the left hand to select for both the mouse and the keyboard conditions. Selection for a mouse is typically done with clicking. Follow-up studies could explore the use of clicking versus keyboard selection. One aspect of clicking is properly holding the mouse. In the mouse training we required participants to show us how they held the mouse, in case they were not familiar with the device. While each passed this step, during the experiment some participants began holding the mouse more on the sides than on the top, gripping with their thumb and pinky finger. This may have been done to improve right-left positioning.

4.1.3.3 Conclusion

This study analyzed the movement times of visually impaired participants using of non-speech audio for a 1-dimensional target selection task. The Fitts's Law index of difficulty had a linear relationship with movement time for both the mouse and the keyboard. Blind and low vision participants were faster with the mouse than with the keyboard. This work also demonstrated the use of key level count as a predictor of movement time.

Blind adults can use mice for 1-dimensional target selection tasks. Between the most common desktop input devices, visually impaired participants select targets with the keyboard 4.2 seconds (49%) slower than with the mouse.

The mouse and the keyboard both appear to be useful input devices for graphing, regardless of level of vision impairment. The use of these devices with auditory displays could benefit those who cannot see a graph. Future work could explore the application of these results to number lines and graphs.

4.2 Study 2: Determine the effect of input device and presentation format on target speed for sighted participants.

A person's level of vision impairment varies widely. Many people are sighted, some are low vision, and a few are blind. Even someone who may be considered by others as completely blind, such as someone who requires a screen reader and a guide dog, often has some basic residual vision. Therefore, providing a visual component to the graphing display alongside the auditory display may help sighted, low vision, and blind students who are graphing. In addition, it allows sighted teachers and others to look at the computer screen and review a student's actions.

The goal of this second study was to identify the average speed for sighted people, using auditory, visual, and combined conditions. In addition, the interaction device was also evaluated (keyboard and mouse). Like Study 1, this is a Fitts's Law experiment. The same expectations for target width and target distance apply: a larger, closer target takes less time to select. Based on [?] and Study 1, people should be faster targeting with the mouse over the keyboard. It was expected that the visual display would be faster than the auditory display. It was not clear whether the combined condition would be much different than the visual condition.

There were the following questions for Study 2:

Q_1 What is the effect of target size on keyboard movement time?

Q_2 What is the effect of target size on mouse movement time?

Q_3 What is the effect of input device on movement time?

Q_4 What is the effect of display format on movement time?

The Fitts's Law questions constitute a single family of statistical tests, and alpha should be adjusted accordingly. The display format tests are independent and were be tested on $\alpha = 0.05$.

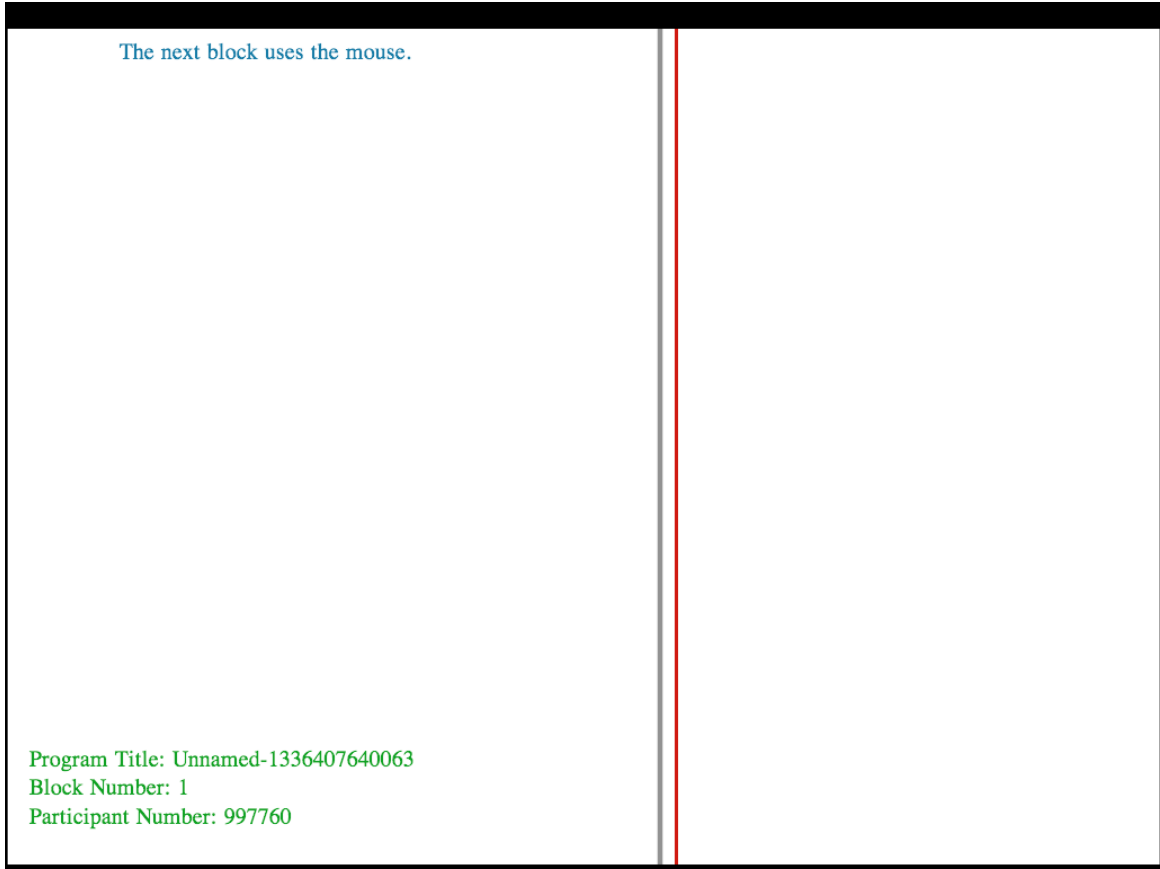


Figure 34: The visuals for Study 2. The vertical gray line is the target. The vertical red line is the cursor. The black edges are the sides of the screen.

4.2.1 Study Design

The auditory display, evaluation program, and basic structure of the study are largely the same as Study 1. Only the changes will be discussed below.

4.2.1.1 Apparatus

The experiment space was in the Coon Psychology building at Georgia Tech, in three small offices and a larger room connected to the offices. Participants used the same mice, keyboards, and headphones as the first study. AudioFitts was adapted to have visual and combined modules, but the software was largely the same.

The same keyboard and mouse settings were used as in the first study. This study also used the same steps for keys, 1, 10, and 100 pixels. This study also used the same auditory display, mapping a higher pitch to a closer distance, with the logarithmic equation 5.

The visuals for Study 2 were simple (Figure 34). They showed a target, in gray, the current x location of the cursor, in red, and some text about the study. When the participant moved the mouse or keyboard (depending on the condition), the cursor would move left or right. When the cursor was on the target, the cursor turned blue, and turn back to red when the cursor moved off the target. During a block, each time a participant pressed space bar, the gray target changed location. Since the participants were blindfolded during the audio-only condition, the visuals were kept on, allowing the research personnel to observe the status of the experiment. This was a promising side effect, as one goal of this research is to create a tool that could be used by two people simultaneously, such as one person (a student) listening, and one person (a teacher) looking.

4.2.1.2 Participants

Students from the Georgia Tech Psychology Pool participated in the study. This population informs best-case performance of the graphing (particularly visuals), since the students will generally be young, have high math (and graphing) skills, and a high level of experience with computers. Therefore, this population should have quick movement time and high accuracy. Participants were recruited through Experimentrix, and given 1 point extra credit for a single-session, one-hour study.

All 29 participants were included in the data analysis. 13 were male (44.8%), 27 were right handed (95.1%), and the mean age was 19.6 (minimum 17, maximum 30)¹⁰. All reported at more than 6 years experience with computers and the mouse¹¹; all but one reported more than 6 years experience with the keyboard (the remaining participant selected 4-6 years). Most (15) participants reported using computers 3 to 4 hours daily; all reported using computers at least one hour per day.

It is important not to over-compare the results of this population with the population from the work rehabilitation center, as the age, computer experience, and mathematics education are much different. Based on these three components, I would expect this population

¹⁰Like the first study, ages were calculated based on 2011 minus the year of birth. One participant entered “Virginia” instead of their date of birth.

¹¹One participant reported not using the mouse, but selected “More than 6” years experience using the mouse.

Table 15: Study 2 significance tests conducted. All results are significant, except for visuals-combined.

| Topic | Result | p-value | Effect Size |
|-------------------------|----------------------|--------------|-------------------|
| Target Width | $F(2, 56) = 38.216$ | $p < 0.0071$ | $\mu_p^2 = 0.644$ |
| Target Width, Contrasts | $F(1, 28) = 50.631$ | $p < 0.0071$ | $\mu_p^2 = 0.577$ |
| Input Device | $F(1, 28) = 102.512$ | $p < 0.0071$ | $\mu_p^2 = 0.785$ |
| Sense | $F(2, 56) = 164.225$ | $p < 0.0071$ | $\mu_p^2 = 0.854$ |
| Visuals-Audio | $t(28) = -13.746$ | $p < 0.0071$ | $d = 2.55$ |
| Audio-Combined | $t(28) = 12.373$ | $p < 0.0071$ | $d = 2.30$ |
| Visuals-Combined | $t(28) = -2.124$ | $p = 0.043$ | $d = 0.395$ |

to perform better than the CVI population (and most any population) for the visual and non-visual tasks.

4.2.1.3 Task

Up to three participants at one time started in the larger room, were given informed consent, and a brief introduction. They were then escorted to one of the smaller rooms, one student per room. Like the first study, participants put on headphones, which they wore for all three conditions (audio, visual, combined). Participants wore sleep shades during the audio-only condition, and would slide them on top of their head or take them completely off during the visual and combined conditions.

There are three variables to this experiment: display format (audio, visual, and combined), target width (5, 9, and 13 pixels), and interaction device (mouse and keyboard), for a total of 18 blocks. The order of the target width was randomized by the system for each participant. The order of the interaction device and sensory conditions was counterbalanced for each participant. Participants had an introduction, followed by a training and experiment for one input device, followed by a training and experiment for the other input device. All three display sense conditions were run (counterbalanced) before moving to the second input device. For more details on the task, see Study 1. The overall experiment lasted about an hour.

4.2.2 Results

Based on the large effect sizes found in study 1, it appeared that the results are robust even with a relatively low number of participants. Therefore, the analysis of the second study

Figure 35: Box plots of the movement time, based on the sense and interaction device.

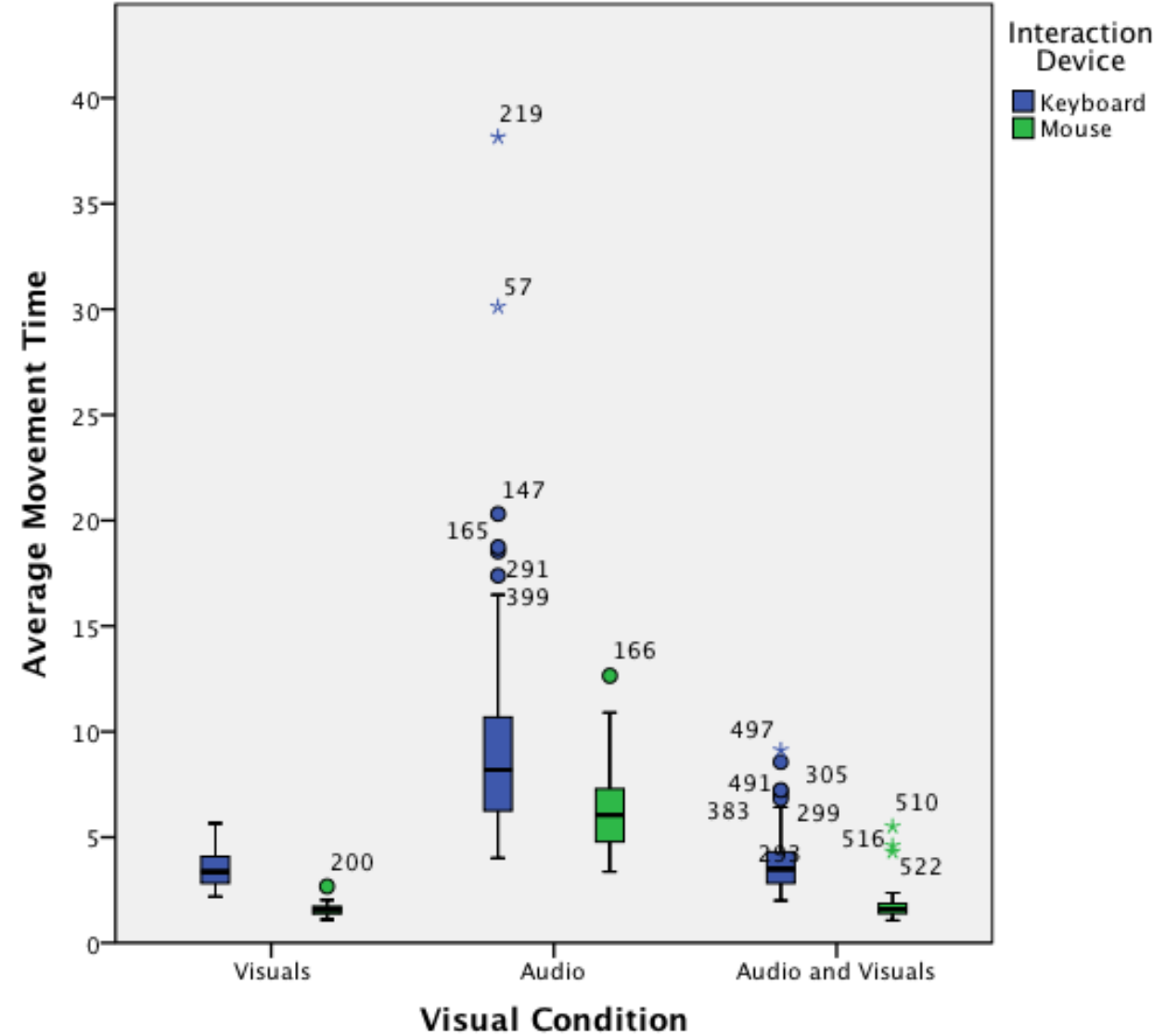


Table 16: Descriptive statistics for speed for the interaction device and sensory condition. The mean values are first, followed by the standard deviation in parentheses.

| Sense | Interaction Device | | |
|-----------------|--------------------|---------------|---------------|
| | Mouse | Keyboard | Overall |
| Visual | 1.570 (0.255) | 3.532 (0.873) | 2.551 (1.174) |
| Audio | 6.205 (1.802) | 9.368 (5.217) | 7.786 (4.203) |
| Combined | 1.735 (0.667) | 3.813 (1.375) | 2.774 (1.499) |
| Overall | 3.170 (2.443) | 5.571 (4.139) | 4.370 (3.595) |

used a Bonferonni correction to adjust the alpha. There were 7 tests, so the adjusted alpha was $\alpha = 0.05/7 \approx .0071$. There are three variables, and so there are many combinations of tests possible. The focus of the tests was on main effects. These were found in all tests, except for the comparison between the visuals and combination condition. See table 15 and Figure 35.

Not surprisingly, movement time depends on target width. As target width increased, movement time decreased, from 5.3 seconds at 5 pixels, to 4.1 seconds at 9 pixels, to 3.7 seconds at 13 pixels. Based on the effect sizes (in Table 15), it is clear that the interaction device and the sense have an even larger effect on movement time.

Accuracy for each target width, interaction device, and sense was over 95% on average, and each had a median of 100%. Fitts's Law studies often seek to have maximum accuracy, and focus on analysis on speed. Study 3 and 4, examining auditory display, look at accuracy in more detail.

The average speed was 4.370 seconds (3.595 seconds). The mean movement time for the keyboard was 5.570 seconds (SD=4.139 seconds). The mean movement time for the mouse was 3.170 seconds (SD=2.443 seconds). The mouse is significantly faster than the keyboard.

The mean movement time using visuals was 2.551 seconds (SD=1.174 seconds). The mean movement time for using audio was 7.786 seconds (SD=4.203 seconds). The mean movement time for using the combination of visuals and audio was 2.774 seconds (SD=1.499 seconds). Visuals and audio are significantly different, as are audio and combination. Visuals and combination are not significantly different, although the combination has a slower mean average for each interaction device, by about 0.2 seconds.

The interaction device and the sense both have large effects on movement time. Table 16 shows descriptives of how the combination of sense and interaction device affect movement time.

4.2.3 Discussion

Based on the results of study , there are the following conclusions. For sighted participants using AudioFitts:

1. As target width decreases, movement time increases.
2. The mouse is generally faster than the keyboard.
3. Having visuals on, either with audio or without, is faster than only having audio on.

Participants again outperformed with the mouse over the keyboard. This is not surprising, given previous Fitts's Law studies [? ?]. One important difference between study 1 and 2 is the participant pool. When only considering the auditory display, the mouse mean was 7.560 seconds in study 1, and 6.205 seconds in study 2, a difference of 1.355 seconds (18%). The keyboard mean was 10.030 seconds in study 1, and 9.368 seconds in study 2, a difference of 0.662 seconds (6.6%). Not surprisingly, young, sighted, math-skilled students appear to be faster than middle-aged, visually impaired, standard-education adults. Mouse performance was proportionally better with the sighted students, possibly due to the higher levels of experience using the mouse.

Participants were faster with visuals only than with audio only. Surprisingly, participants were slightly slower with the combined condition (not significantly, but the means difference was about 0.2 seconds slower on average). Perhaps the cognitive load of visual and auditory information was higher, or the information somehow conflicted, so participants had to spend more time understanding the display. This is an interesting avenue of further research, but if there is an effect, it is relatively small, and will not be further explored in this dissertation.

Based on study 2, it is clear that visuals (and combinations) are faster than only using audio. For the mouse, it is 3.95 times faster to use visuals, for the keyboard it is 2.65 times

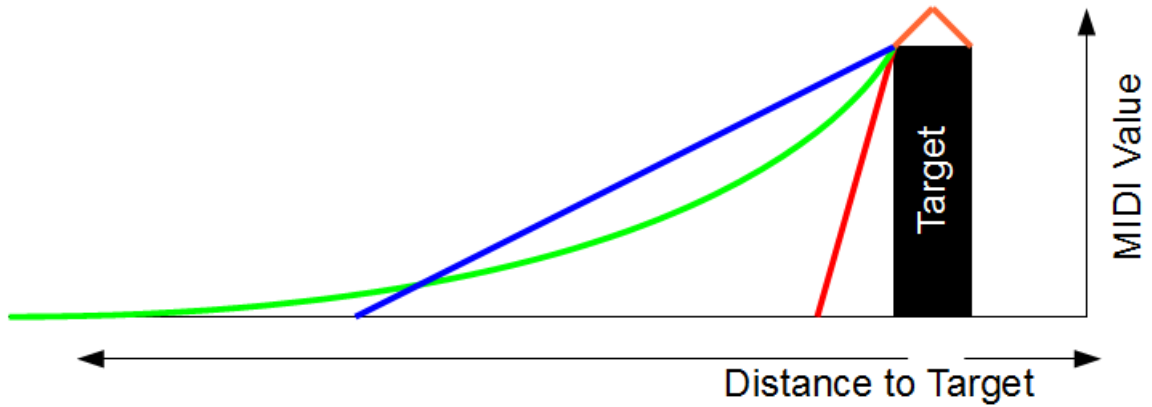


Figure 36: A diagram showing different scalings and peaks. Boolean scaling is vertical, following the black box down. LinearShort scaling is red, dropping quickly but a change for each pixel distance. LinearLong is blue, with a MIDI note drop every 3 pixels and with changes 3 times further than LinearShort. Logarithmic, green, drops the most at first, but has small changes hundreds of pixels away. Peaking is shown in orange. If peaking is on, MIDI notes change when the user’s selection pixel is within the target; if off, the MIDI note does not change.

faster, and overall, it is 3.05 times faster. In many cases, the target user group for the auditory graphs will not be able to use the visual components. However, the ratio of time of visuals to audio is informative for testing accommodations. A first guess at the time it takes for a blind student to answer a graphing question may be about three times the amount of time given to a sighted (or low vision) student. This will be further evaluated in phase 3, studies 1 and 2, where sighted and visually impaired people answer test questions.

Studies 1 and 2 show the effect of target width, target distance range, interaction device, and sense on the movement time. There are still questions about the basic auditory display to use for point estimation.

4.3 Study 3: Determine the effect of auditory scaling and peaking on target speed and accuracy.

Study 3 explored two aspects of the auditory display: “auditory scaling” and “peaking”. One factor that may affect speed and accuracy is the way that the distance mapping is transformed to pitch. Different auditory scalings may make it easier to find the target. Four scalings were evaluated. A “Logarithmic” scaling has more change close to the target, as described in M_1 . A “Boolean” scaling has a low pitch when off the target, and high

Table 17: The different scalings evaluated. The top row is the distance from the target, and the following rows show the MIDI note for each scaling. The actual MIDI range will be about 40-80. For the formulas, the MIDI note is bounded within the range.

| Scaling | Formula | Distance from target | | | | | | | | | | |
|--------------------|----------------------------------|----------------------|----|----|----|----|----|----|----|----|----|-----|
| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 10 | 20 | 50 | 100 |
| Boolean | $N = 80 - (40 * d)$ | 80 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 | 40 |
| LinearShort | $N = 80 - d$ | 80 | 79 | 78 | 77 | 76 | 75 | 74 | 70 | 60 | 40 | 40 |
| LinearLong | $N = \lfloor 80 - (d/3) \rfloor$ | 80 | 79 | 79 | 79 | 78 | 78 | 78 | 76 | 73 | 63 | 46 |
| Logarithmic | (Equation 5) | 80 | 75 | 73 | 71 | 70 | 69 | 68 | 66 | 62 | 57 | 53 |

pitch when on the target, with no intermediate changes. A “Linear Short” scaling changes in pitch by one MIDI note for each pixel distance from the target, until the full range of notes is used. This, however, runs out within about 30 pixels, which may not be sufficiently informative. A “Linear Long” scaling changes 1 MIDI note for every 3 pixels, increasing the range of information around the target.

In studies 1 and 2, participants who find a target might only move to the edge of the target and select. A miniscule movement of the mouse could shift their targeted pixel off the edge of the intended target. To rectify this, a “peak” display was used to show how close to the center of the target the user is, while still indicating that the user is within or outside of the target. The peak portion always used a “Linear Short” scaling, increasing for each point toward the center of the target. The rate and pitch warble between octaves remained, indicating that the user was within the target. The peak options are “peak” or “flat”, and were applied to all of the audio scaling conditions. One possible outcome for peak was that participants would be slower, but more accurate, since they might spend more time finding the center. See Figure 36.

There were the following questions for study 3:

Q_1 What is the effect of auditory scaling on target speed?

Q_2 What is the effect of auditory scaling on target accuracy?

Q_1 What is the effect of peaking on target speed?

Q_2 What is the effect of peaking on target accuracy?

4.3.1 Study Design

The targeting task used the same AudioFitts evaluation tool as studies 1 and 2. Other than the use of a number line, described below, and the new auditory scaling and peaking, the experiment was very similar to the first two studies.

4.3.1.1 Apparatus

The experiment space was a large room at CVI. Laptops were purchased for the study, and brought to the study location. The computers were ASUS Aspire 5750Z-4877 laptops, with a 2.0 GHz processor, 4GB memory, and a 15.6" screen. The Operating System was Windows 7 Professional 64-bit. Participants wore Sennheizer HD 202 headphones and sleep shades. The same mouse settings were used as in study 1. Participants only used the mouse for this study.

In studies 1 and 2, there was a single target. In this study, the system displayed several targets, representing tick marks on a number line. The tick marks had a target width of 9 pixels. In addition, each tick mark had a label, with numbers within the range of 1 to 30¹², and always included the number 15. Tick marks had a step size of 1, with numbers increasing from left to right. See Figure 37.

For the auditory display, the distance to the target was mapped pitch, with the peak and specific audio scaling based on the condition for the block. The system selected the closest tick mark to measure distance. If the user had the cursor of the number line (past the arrows), the system played a brown noise.

¹²Most adults should be able to understand whole numbers on a number line from 1 to 30. GPS Kindergarten MKN1 [?] expects students to

- a. Count a number of objects up to 30.
- b. Produce models for number words through ten.
- c. Write numerals through 20 to label sets.
- d. Sequence and identify using ordinal numbers (1st-10th).
- e. Compare two or more sets of objects (1-10) and identify which set is equal to, more than, or less than the other.

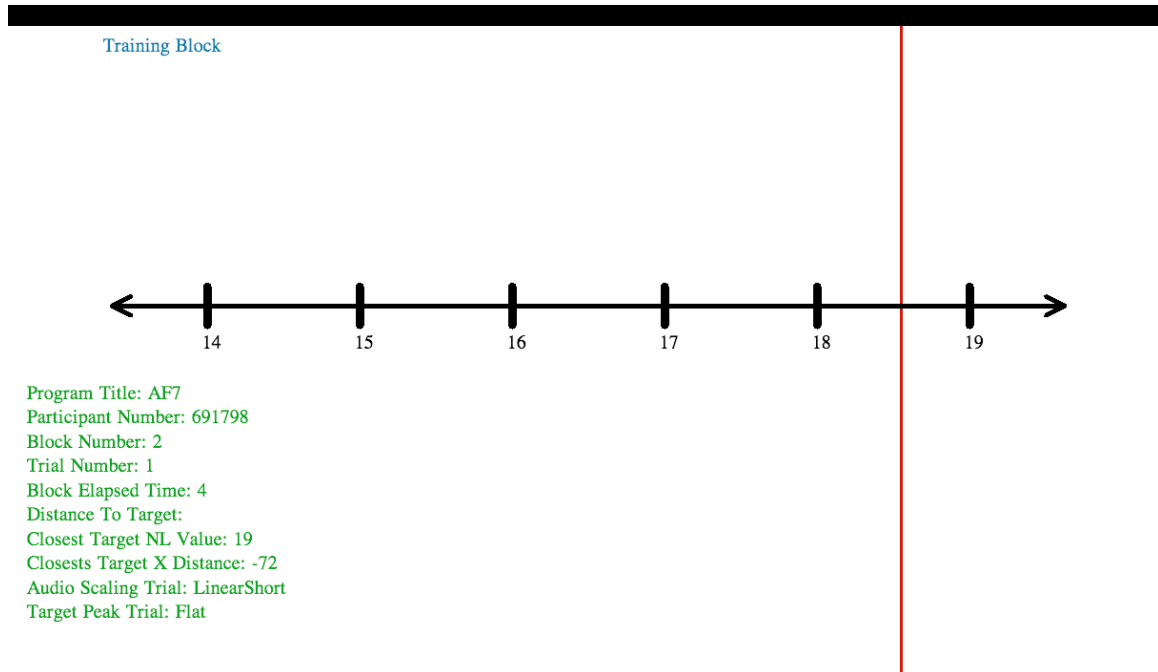


Figure 37: The number line used in study 3.

4.3.1.2 Participants

Visually impaired participants were recruited from the Center for the Visually Impaired (CVI). The participants were blindfolded. Participants were compensated \$15 for their time.

Many of the participants had completed similar studies¹³, including study 1. However, participation in the earlier studies will probably not affect within-subjects results. Based on the limited supply of the visually impaired adult population, studies 3 and 4 reuse some participants from previous studies.

4.3.1.3 Task

For each experiment session, 1 to 4 participants were given information about the study and signed informed consent documents with a reader and a witness. Participants were then given a brief introduction to number lines. A braille-tactile graphic of a number line and a enlarged number line was presented, with the numbers 1, 2, and 3. Participants were asked

¹³There are four other studies in the AudioFitts line of research which does not appear here that participants may also have completed.

Table 18: Accuracy of peak and auditory scaling conditions for Study 3.

| Audio Scaling | Peaking | | |
|--------------------|---------------|---------------|---------------|
| | Flat | Peak | Overall |
| Boolean | 0.875 (0.156) | 0.919 (0.144) | 0.897 (0.150) |
| LinearShort | 0.877 (0.147) | 0.943 (0.071) | 0.910 (0.119) |
| LinearLong | 0.931 (0.091) | 0.905 (0.150) | 0.918 (0.123) |
| Logarithmic | 0.945 (0.071) | 0.938 (0.097) | 0.941 (0.084) |
| Overall | 0.907 (0.124) | 0.926 (0.119) | 0.917 (0.122) |

to find the number 3. People could use their vision during this demonstration. Participants repeated this task with a second number line, finding the number 3 with the numbers 2, 3, and 4. We then explained the experiment task is the same, except the target number will always be 15, the number line is on a computer, and participants would be hearing the number line instead of seeing it.

Participants were then led to one of the four laptop computers. They were presented with a brief, interactive training on finding the number 15 on a number line. Participants were then presented with a training about the audio scaling and peaking. During this training block, participants received a rotating combination of peaks and audio scalings. Selection involved finding with the mouse (moving left and right) and clicking with the mouse button. Participants had to select the number 15 correctly eight times in a row to complete the block; incorrect selections reset the hits-in-a-row counter to 0. During the training, when participants selected correctly, the system said “hit”, and when participants selected incorrectly, the system said “miss.” Participants reacted strongly to this verbal feedback, and it improved their hit rate.

Experiment blocks lasted 180 seconds. With four audio scaling conditions and two peaking conditions, there were eight blocks. These blocks were randomized by the system for each participant. At the end of the experiment, participants answered a survey.

4.3.2 Results

For this study (3) and study 4, the purpose is to generally understand the effects of the audio scaling and the peaking on speed and accuracy. While it remains a reaction time experiment, it is not a Fitts’ Law study, like study 1 and 2. Instead, the purpose is to make

Table 19: Speed of peak and auditory scaling conditions for Study 3. In seconds.

| Audio Scaling | Peaking | | |
|--------------------|----------------|-----------------|----------------|
| | Flat | Peak | Overall |
| Boolean | 11.249 (4.181) | 13.422 (4.854) | 12.335 (4.618) |
| LinearShort | 11.744 (4.910) | 12.922 (6.854) | 12.333 (5.933) |
| LinearLong | 12.813 (5.508) | 15.308 (12.217) | 14.061 (9.467) |
| Logarithmic | 11.886 (5.065) | 13.349 (6.429) | 12.617 (5.777) |
| Overall | 11.923 (4.900) | 13.750 (8.015) | 12.837 (6.689) |

design decisions about the use of peaking or audio scaling. If some choices have particularly low accuracy, or high movement time, they probably are not a good choice for the design.

Neither audio scaling nor peak had a significant main effect for either hitrate or time. For audio scaling hitrate, $F(3,75)=2.078$. For audio scaling time, $F(3,75)=3.099$. For peak hitrate, $F(1,25)=2.714$. For peak time, $F(1,25)=2.414$. Since there was no main effect, individual effects were not compared. The statistics below and in Tables 18 and 19 are descriptive, and do not indicate significant differences.

Having a target peaking increased accuracy from 90.7% to 92.6%, but slowed down performance from 11.9 seconds to 13.8 seconds (see Tables 18 and 19). Having more information about where the user is in within the target appears to increase the user’s chance of hitting the target. It is likely that the additional targeting time is due to the extra time the user is taking to find the target.

For audio scaling, accuracy averaged 91.7%. The lowest performance was from the Boolean display (89.7%). The Boolean display also was tied for the fastest time (12.335 seconds, next to 12.333 seconds for LinearShort). This is somewhat surprising, since in pilots, it felt as if it was harder to find the target without additional the auditory feedback; it looks as if this “location confidence” doesn’t manifest in slower speeds, however. The highest accuracy was with Logarithmic (94.1%), but this may have been due to experience with Logarithmic on previous experiments.

4.3.3 Discussion

Based on the results of this study, there are the following conclusions.

1. The overall speed is about 12.8 seconds, with a range of 12.3-14.1 scaling for the audio

scaling conditions and 11.9-13.8 seconds for the peak condition.

2. The overall accuracy is about 91.7%, with a range of 89.7%-94.1% for the audio scaling conditions and 90.7%-92.6% for the peak condition.
3. The differences in the speed and accuracy between conditions are not statistically significant for 25 visually impaired participants.

Study 3 did not find conclusive difference between the different audio scalings or peak. This suggests that, while there may be differences, the effect is not large. People can map pitch to the distance from the target with a variety of stimuli, suggesting that the concept is robust for different display choices.

It is surprising that participants performed so well with the Boolean condition. It is tempting to state that the Boolean condition shows that no auditory cues are necessary when close to a target. However, the nature of this task was to find a target; participants knew there would be a 15 on every trial. In graphing situations, a point may or may not be present. Anecdotally, blind people have commented on the “inch is a mile” phenomenon, when searching for something and missing it regardless of how close. Therefore, this proximal information could be more useful when scanning a graph when point locations are not known.

Distance information may be less necessary for predictable situations, such as presenting tick marks along a regular interval. Imagine a student moving the mouse at a constant speed from left to right over a number line. The graph system speaks “3”, “4”, “5” at a regular time interval. The student can expect the next number to be “6”, and can probably guess the distance and time to approach the “6”. However, the student has no idea where a point could be; it doesn’t need to be near a tick mark, or at regular intervals from other points. In the tick mark case, a Boolean mapping could be used. In the point case, it will probably be worth the information processing overhead to have distance feedback.

While in an ideal situation the distance would always be provided, it may simply overload the auditory information. If context sounds such as x-axis y-axis locations, and value sounds such as the nearest point are played simultaneously, the auditory display will present three

simultaneous tracks of sound. To reduce the cognitive workload of understanding the sound, the predictable markers could use a boolean audio scaling.

4.4 Study 4: Determine the effect of mapping type and auditory scaling on target speed and accuracy.

The final study in this phase included the type of auditory mapping used. There were four types: pan, pitch, rate, and volume. Only pitch had been evaluated in the first three studies.

While pitch appears to be the best mapping (e.g. [?]), the other types may be suitable for point estimation. In addition, pitch may already be used for a certain mapping. The final interactive display may be providing several pieces of information to the user; a sole reliance on pitch may reduce the quality of the interface.

4.4.1 Study Design

The types of auditory display that were mapped to the distance are pitch, pan, rate, and volume. Pitch works as before, increasing when moving toward the target; if using a different display type, pitch stays at a fixed MIDI value.

Pan is mapped to left-center-right, where a distance of 0 is at the center. Unlike the other formats, pan includes bimodal information about being to the left or the right of the target. If using a different display type, the pan remains at center. While using left/right may be initially appealing, one consequence is a highly noticeable artifact equidistant between two tick marks, where the pan changes suddenly from right to left (and vice versa).

The “rate” is the repeat speed of new notes played in MIDI. Imagine a person pressing a piano key; the pitch stays the same, but the attack on the key is louder than the following sound. This rate of new notes is fast when at the target, and slow when away from the target. If using a different display type, rate remains slow.

The volume mapping is louder when at the target, quieter when away from the target. If using a different display type, the volume remains fixed. All display types were mapped according to the audio scalings conditions of Boolean, Linear Long, and Logarithmic, described in study 3. See Table 17.

Table 20: Study 3 results.

| Audio Type | Audio Scaling | Accuracy | significance | Effect Size |
|------------|---------------|-----------------|-------------------|-------------|
| Pan | Boolean | 0.6776 (0.3660) | $t(20) = -2.785$ | - |
| | LinearLong | 0.1961 (0.2188) | $t(20) = -14.741$ | $d = 3.22$ |
| | Logarithmic | 0.4864 (0.2914) | $t(20) = -6.505$ | $d = 1.42$ |
| Pitch | Boolean | 0.9443 (0.0826) | $t(20) = 2.457$ | - |
| | LinearLong | 0.5578 (0.3739) | $t(20) = -4.193$ | $d = 0.915$ |
| | Logarithmic | 0.8918 (0.2120) | $t(20) = -0.176$ | - |
| Rate | Boolean | 0.6195 (0.3898) | $t(20) = -3.297$ | - |
| | LinearLong | 0.3642 (0.3476) | $t(20) = -7.064$ | $d = 1.54$ |
| | Logarithmic | 0.7364 (0.3419) | $t(20) = -2.193$ | - |
| Volume | Boolean | 0.9254 (0.2097) | $t(20) = 0.555$ | - |
| | LinearLong | 0.3058 (0.2359) | $t(20) = -11.542$ | $d = 2.52$ |
| | Logarithmic | 0.8519 (0.2399) | $t(20) = -0.919$ | - |

This study used the same equipment as study 3, and the same room at CVI. Also like study 3, this study used a number line, with the goal to find the target 15. Unlike study 3, participants had to press spacebar instead of clicking the mouse to select the target.

Visually impaired participants were recruited from the Center for the Visually Impaired (CVI). The participants were blindfolded. Participants were compensated \$15 for their time.

The task was very similar as study 3, with a few exceptions. Instead of 6 conditions at 180 seconds, there were 12 conditions at 105 seconds each. In addition, the training presented the pan, pitch, rate, and volume sections as independent trainings, where the participant had to attempt 5 trials before moving to the next sound type. Finally, the audio scalings were reduced to Boolean, LinearLong, and Logarithmic.

4.4.2 Results

(need the speed table, and probably an overall accuracy table.)

For a more complete picture, the data is displayed with both independent variables in table 20, and separately in tables *(make tables for speed and accuracy, for scalings and type)*. Differences were more apparent than in study 3.

Table 20 shows a test of whether the accuracy for the condition is significantly less than 90%. This test differentiates good candidates for targeting from those that are much more difficult. From this test, the LinearLong condition for each sound type was significantly

slower than 80%. In addition, the Logarithmic-Pan condition was also slower.

4.4.3 Discussion

In terms of descriptives, pitch and volume outperformed rate, which outperformed volume. The most accurate targeting using sound types, in order from greatest to least, are pitch, volume, rate, and pan. The fastest targeting using sound types, in order from fastest to slowest, are volume, pitch, rate, and pan.

For audio scalings, Boolean and Logarithmic scalings did particularly well, while Linear-Long scalings produced average accuracies 30-40% lower. This is surprising, since in study 3 (pitch only), LinearLong performed about as well as Logarithmic, and often better than Boolean. The variance in pitch-LinearLong is high, however, and the other sound types may not be as robust as pitch in displaying targeting in different scalings. Nevertheless, the use of LinearLong should be used cautiously.

It appears that people can use different types of sounds for finding a target. However, the system is not as robust. Depending on the condition, accuracy can be have a high average of 90% or a low of 20%. Many people in the low-performing conditions had 0% correct.

Later tests should be conducted to evaluate the settings within the situation. Redundant information in the form of combinations of conditions, such as Rate-Logarithmic and Pitch-Linear, may lead to better (or worse) speed and accuracy over the original conditions.

4.5 General Discussion

Based on the phase 1 studies, it appears that sighted, low vision, and blind people can use a mapping of pitch to distance to find a target. Participants could find the target using a mouse or a keyboard. They had varying levels of success, depending on the audio scaling, peak, and sound type. This work extends the trend studies started by Mansur and Blattner [?], and could be used to enhance previously proposed systems such as the IC2D [?], MULTIVIS [?], Sonification Sandbox [? ?], and AHEAD [?].

These experimental studies are informative, but do not examine more realistic scenarios. On a real number line or graph, it may be necessary to simultaneously display multiple

pieces of information. For example, tick marks and labels may need to be presented at the same time as the distance to a point. These sounds must be managed so that they do not become overwhelming to the user. This more realistic development and evaluation will be conducted in phases 2 and 3.

CHAPTER V

PHASE 2: FROM GRAPHING STANDARDS TO A GRAPHING SYSTEM

With 30 years of trend analysis research, bolstered by the studies in point estimation described in phase 1, there is sufficient basic research to be able to build a system. As mentioned in the Background (Chapter 2), many other systems have been built to cover certain parts of graphing. However, the primary challenge remains: creating an accessible graphing system that can reasonably be expected to cover the graphing requirements for a particular education course.

A new method, Standards, Questions, Answers, Reconstruct, and Evaluate (SQUARE), was used for this dissertation. SQUARE has the following steps:

1. Standards. Identify education standards used in graphing.
2. Questions. Find and create graphing questions based on the standards.
3. Answers. Generate a task analysis based on answering the graphing questions.
4. Reconstruct. Build a system that can be used to answer the same questions using the same steps.
5. Evaluate. Confirm that the new system can be used to answer the same questions using the same steps.

5.1 Standards: Identify education standards used in graphing.

Standards identification is a classification process. A list of standards are categorized as either being a graphing standard or not a graphing standard. In practice, a single standard may have a complex description. For example, it may be difficult to create a single, simple problem to comprehensively cover the entire standard. In these cases, a process for breaking down the standard into pieces was necessary. In the end, the “graphing standards” are

adaptations of the original standards, that use the same text as the original standard but might not always include the complete text of the entire standard.

Starting with the standard has a distinct advantage over basing development on a sample of graphing questions. The standards are the learning expectations for the course; a student demonstrating her completion of each standard should lead to her passing the course. A sample of questions only demonstrates that the student can complete those questions. Other questions that are important to the course may not be solvable with the tool. The exception, of course, is if those questions are samples of each standard. Then, taken together, answering the questions correctly is the demonstration that the student understands the standards.

5.1.1 Study Design

The standards used in this study were the Common Core Standards for Mathematics, Grade 6 (available at [?] and as appendix A). Sixth grade is the midpoint of K-12 education, and has many graph and number line standards. This grade will give an indication whether this process and auditory graphs could be used in education, without exclusively focusing on the easiest or the hardest questions in primary and secondary school.

Participants were three researchers from the Georgia Tech Sonification Lab. One researcher (the author) identified standards that included graphing, and these were adapted to standards with a single objective. The other two researchers assisted with the adaptation.

Standards were found through a word search. A list of words related to graphing were used to find the standards containing those key words. The words for this search were “coordinate”, “graph”, and “number line”.

After the initial pass, these standards were filtered to only include standards that include a one-dimensional number line and a two-dimensional coordinate graph. Other forms of graphs, such as double number lines and dot plots, were excluded for the purposes of this investigation. Other graphs appear to

Finally, researchers considered whether the standard could be satisfied with a single graphing question. If not, it was broken down into pieces. The single-answer graphing standards were then re-labeled with the new graph name.

5.1.2 Results

In part 1, 12 standards had the words “coordinate”, “graph”, or “number line”. Of the standards found in part 1, two did not appear to use plain number lines or coordinate graphs, and were excluded in part 2. There are 47 standards, so the 10 remaining represent 21% of the standards for sixth grade math. See Table 21.

For part 3, five standards were split into two or three new graphing standards. For example, 6.NS.6 had the following text:

Understand a rational number as a point on the number line. Extend number line diagrams and coordinate axes familiar from previous grades to represent points on the line and in the plane with negative number coordinates.

The first sentence of this standard discusses rational numbers and number lines, and was selected as the first graphing standard. The second sentence discusses both number lines and graphs. The second and third graphing standards selected text to describe a number line and a coordinate graph that would have negative number coordinates. As one standard, it would be difficult to evaluate in a single, simple graphing question. Split apart, however, the three pieces sufficiently represent the graphing portion of the standard. See Table 22.

5.1.3 Discussion

The sixth grade Common Core Standards have 13 standards related to graphing, which have been adapted into 17 graphing standards. Seven of the standards use number lines, and eleven use graphs. The importance of graphing is obvious given that 21% of the standards include number lines and graphs. A student who does not have the tools to learn graphing can, at best, demonstrate competence in 79% of the standards.

The Common Core Standards are heirarchically arranged, for example 6.NS.6.a is under 6.NS.6. However, below the section indicator (e.g. “NS” for “Number Sense”), there is no clear level of importance. A branch contains a tree, but it does not appear to be true that all branches are equally important. To fully represent the content of the standards, all levels of the standards were considered. Depending on the intention of the standards, this may

Table 21: Standards results, parts 1 and 2: The graphing standards that include the text “coordinate”, “graph”, and “number line”. Standards excluded in part 2 have a strikethrough.

| ID | Standard |
|-------------------|--|
| 6.RP.3 | Use ratio and rate reasoning to solve real-world and mathematical problems, e.g., by reasoning about tables of equivalent ratios, tape diagrams, double number line diagrams, or equations. |
| 6.RP.3.a | Make tables of equivalent ratios relating quantities with whole-number measurements, find missing values in the tables, and plot the pairs of values on the coordinate plane. Use tables to compare ratios. |
| 6.NS.6 | Understand a rational number as a point on the number line. Extend number line diagrams and coordinate axes familiar from previous grades to represent points on the line and in the plane with negative number coordinates. |
| 6.NS.6.a | Recognize opposite signs of numbers as indicating locations on opposite sides of 0 on the number line; recognize that the opposite of the opposite of a number is the number itself. |
| 6.NS.6.b | Understand signs of numbers in ordered pairs as indicating locations in quadrants of the coordinate plane; recognize that when two ordered pairs differ only by signs, the locations of the points are related by reflections across one or both axes. |
| 6.NS.6.c | Find and position integers and other rational numbers on a horizontal or vertical number line diagram; find and position pairs of integers and other rational numbers on a coordinate plane. |
| 6.NS.7.a | Interpret statements of inequality as statements about the relative position of two numbers on a number line diagram. |
| 6.NS.7.c | Understand the absolute value of a rational number as its distance from 0 on the number line; interpret absolute value as magnitude for a positive or negative quantity in a real-world situation. |
| 6.NS.8 | Solve real-world and mathematical problems by graphing points in all four quadrants of the coordinate plane. Include use of coordinates and absolute value to find distances between points with the same first coordinate or the same second coordinate. |
| 6.EE.8 | Write an inequality of the form $x > c$ or $x < c$ to represent a constraint or condition in a real-world or mathematical problem. Recognize that inequalities of the form $x > c$ or $x < c$ have infinitely many solutions; represent solutions of such inequalities on number line diagrams. |
| 6.EE.9 | Use variables to represent two quantities in a real-world problem that change in relationship to one another; write an equation to express one quantity, thought of as the dependent variable, in terms of the other quantity, thought of as the independent variable. Analyze the relationship between the dependent and independent variables using graphs and tables, and relate these to the equation. |
| 6.G.3 | Draw polygons in the coordinate plane given coordinates for the vertices; use coordinates to find the length of a side joining points with the same first coordinate or the same second coordinate. Apply these techniques in the context of solving real-world and mathematical problems. |
| 6.SP.4 | Display numerical data in plots on a number line, including dot plots, histograms, and box plots. |

Table 22: Standards results, part 3: The final graphing standards, based on CCS Mathematics Grade 6.

| ID | CCS | Type | Text |
|-------|----------|-------|---|
| GS.1 | 6.RP.3.a | Graph | Make tables of equivalent ratios relating quantities with whole-number measurements, find missing values in the tables, and plot the pairs of values on the coordinate plane. |
| GS.2 | 6.NS.6 | NL | Understand a rational number as a point on the number line. |
| GS.3 | 6.NS.6 | NL | Extend number line diagrams familiar from previous grades to represent points on the line with negative number coordinates. |
| GS.4 | 6.NS.6 | Graph | Extend coordinate axes familiar from previous grades to represent points in the plane with negative number coordinates. |
| GS.5 | 6.NS.6.a | NL | Recognize opposite signs of numbers as indicating locations on opposite sides of 0 on the number line. |
| GS.6 | 6.NS.6.b | Graph | Understand signs of numbers in ordered pairs as indicating locations in quadrants of the coordinate plane. |
| GS.7 | 6.NS.6.b | Graph | recognize that when two ordered pairs differ only by signs, the locations of the points are related by reflections across one or both axes. |
| GS.8 | 6.NS.6.c | NL | Find and position integers and other rational numbers on a horizontal number line diagram. |
| GS.9 | 6.NS.6.c | Graph | Find and position pairs of integers and other rational numbers on a coordinate plane. |
| GS.10 | 6.NS.7.a | NL | Interpret statements of inequality as statements about the relative position of two numbers on a number line diagram. |
| GS.11 | 6.NS.7.c | NL | Understand the absolute value of a rational number as its distance from 0 on the number line. |
| GS.12 | 6.NS.8 | Graph | Solve real-world and mathematical problems by graphing points in all four quadrants of the coordinate plane. |
| GS.13 | 6.NS.8 | Graph | Include use of coordinates and absolute value to find distances between points with the same first coordinate or the same second coordinate. |
| GS.14 | 6.EE.8 | NL | Write an inequality of the form $x > c$ or $x < c$ to represent a constraint or condition in a real-world or mathematical problem. Recognize that inequalities of the form $x > c$ or $x < c$ have infinitely many solutions; represent solutions of such inequalities on number line diagrams. |
| GS.15 | 6.EE.9 | Graph | Write an equation to express one quantity, thought of as the dependent variable, in terms of the other quantity, thought of as the independent variable. Analyze the relationship between the dependent and independent variables using graphs and tables, and relate these to the equation. |
| GS.16 | 6.G.3 | Graph | Draw polygons in the coordinate plane given coordinates for the vertices. |
| GS.17 | 6.G.3 | Graph | Use coordinates to find the length of a side joining points with the same first coordinate or the same second coordinate. |

over-represent or under-represent certain standards in terms of their importance.

The standards identification process involved researchers and may be biased. It is possible that certain standards could be overlooked. A researcher, for example, could intentionally exclude a standards since it would be difficult to implement, and justify the exclusion on the somewhat ambiguous criteria described above. If this becomes a problem, a more rigorous method of validation may be necessary. Content experts (mathematics teachers) could examine a list of standards and validate that certain standards require graphs or number lines. For the purposes of this study, it appeared that the use of graphs and number lines was not controversial so no formal validation was conducted. Having multiple researchers contributed to the development of the final graphing standards, and probably reduced the chances of omission.

5.2 Finding Relevant Questions

Graphing standards show what a student is supposed to know how to do at the end of the course. However, many of the details are not explained in the standard. For example, it may be unclear how detailed a rational number should be, or the typical ranges of numbers for the grade level. To paint a more detailed picture, specific graphing questions are necessary. The main part of this step is question creation, followed by a quick validation of the questions.

Question creation identifies the sort of graphing questions asked in the classroom and on exams. Text resources to answer this question include state and national standards, exam questions, and textbooks (Figure 38). The primary learning requirements for every K-12 grade in Georgia is described in the Georgia Performance Standards (GPS)¹. Students demonstrate proficiency in the GPS objectives in class and ultimately on a test, such as the Mathematics 1 End of Course Tests (EOCTs). While useful resources, the GPS and the EOCT unfortunately do not have sufficient examples and detail for analyzing the broad scope of graphing problems that will have to be made accessible. The standards are too vague, and the EOCT is not sufficiently comprehensive.

In this process, a diverse collection of standards-aligned graph questions are required.

¹Georgia is transitioning from GPS to include CCS between 2012-2015. The combined standards will be called Common Core Georgia Performance Standards (CCGPS), which, must be at least 85% CCS.

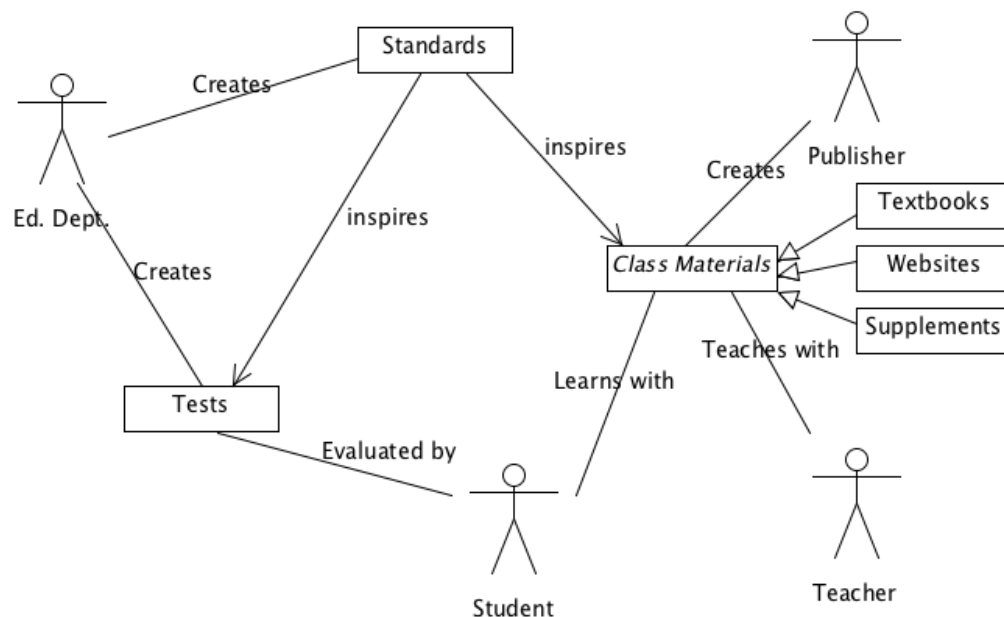


Figure 38: A diagram of how the standards are pushed into the examination and the textbook. Both the tests and the textbook are inspired by the standards. However, the textbook has a different creator (publishers), and a particular graphing question may be less aligned with the standards than most graphing questions on a test.

Textbooks and supplementary texts are a useful starting point. They provide many examples, and many relate problems to specific learning objectives. For example, the Mathematics 1 textbook for Georgia [?] lists specific objectives at the start of each section. A single textbook supplies hundreds of graphing problems, giving a sufficient variety of problems for systems developers to understand how the graphs are being displayed and the specific types of questions.

Textbooks, however, are not necessarily in line with the state standards (Figure ??). In Georgia, standards and tests are created by the Department of Education, while textbooks are created by third-party publishers². While textbooks may be a useful choice for problem selection, there should be an evaluation of how well these questions and other match the

²The Common Core Standards will have a separate standards body from the test publishers. However, since the tests are designed to demonstrate standards competency, and since there are a lot fewer questions, it is likely that the quality of these questions will be higher than a typical textbook question.

standards. *Validated questions* will then provide a pool of problems a developer can confidently use to better understand what a system needs to accomplish. Instead of “teaching to the test” or “teaching to the textbook”, this approach “teaches to the standards”, by finding problems that represent portions of the standards being taught for the class. The tool was designed with the help of the textbook, teacher, and students, leading to a good chance of matching educational needs.

5.2.1 Study Design

Each graphing question is constrained in a few ways.

1. It has a single main idea (no compound questions).
2. It has a specific standard it is intended for, even if other standards may be applicable.
3. There is exactly one graph or number line, a space for text about the graph problem, and a space for writing about the graph.
4. Each answer requires one of three activities by the question solver: answer a multiple choice question based on reading the graph, write out an answer based on reading the graph, or write an answer by editing the graph.

For the types of activities, the first two are called “reading” questions, since they require graph reading literacy. The third activity is called “writing”, since it requires graph writing literacy. Note that reading and writing text outside of the graph may be necessary for all three answering activities.

Many real graphing questions from textbooks and tests may be more complicated than the constrained types used in this study. In Figure 39, only one graphing question could be used immediately. However, in realistic situations, simple questions could be combined to create more complicated questions in realistic scenarios. There may be an issue with the graphing questions used being too simple to represent realistic classroom or test situations. This could become clear during classroom evaluation.

Once a set of all three answer types of graphing questions for each validated graphing standard are available, the graphing questions are validated. Teachers who have experiences

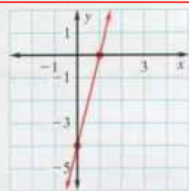
UNIT 1

Exercise Set B *(continued)*

28. Error Analysis Describe and correct the error in finding the intercepts of the line shown.

The line has an x-intercept of -4 and a y-intercept of 1 .

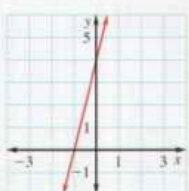
X



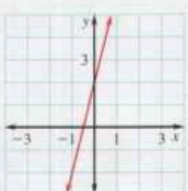
Match the equation with its graph.

| | | |
|--------------------------|---------------------------|---------------------------|
| 29. $8x - 2y = 4$ | 30. $8x - 2y = -4$ | 31. $8x - 2y = -8$ |
|--------------------------|---------------------------|---------------------------|

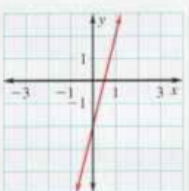
A.



B.



C.



32. Multiple Representations You earn \$16 an hour mowing lawns and \$10 an hour washing windows. You want to make \$500 in one week. This situation can be represented by the equation $16x + 10y = 500$ where x is the number of hours you mow lawns and y is the number of hours you wash windows.

- Graphing an Equation** Find the intercepts of the graph of the equation. Graph the equation.
- Interpreting Intercepts** What do the intercepts mean in this situation?
- Finding Solutions** What are three possible numbers of hours you could work at each job?
- Calculating a Value** If you work 30 hours washing windows, how many hours do you have to mow lawns?

33. Fruit Baskets A small mail-order company that sells fruit baskets currently has 400 orders for fruit baskets. The person who assigns workers to tasks is trying to figure out how many workers need to be assigned to assembling fruit baskets in order to get them out on time. To do this, the person needs to know how many fruit baskets must be produced in one hour. The number B of fruit baskets left to assemble can be modeled by the function $B = 400 - nh$ where n is the number of fruit baskets that can be assembled in one hour and h is the number of hours the company has to produce the fruit baskets.

- Graph the function if the baskets have to be assembled in 48 hours. Then identify the domain and range of this function. How many baskets per hour have to be assembled?
- Suppose 220 baskets are made in the first 24 hours. How does this affect the graph? How many baskets per hour have to be assembled to finish the job?

22

Georgia High School Mathematics 1

Figure 39: Page 22 from the Mathematics 1 textbook [?]. All six problems on this page are graphing problems, and a red box is superimposed around each one. Based on the criteria for graphing question creation, problem 28 is the only one that could be used immediately; all of the other ones would have to be simplified.

with both vision impairment and mathematics were raters. Questions validation involved qualitative coding of graphing questions. Raters were presented with the CCS description of the specific standard, and the graphing question (the text and the graph). Raters determined whether the graphing question either covers or does not cover the standard. Since the graphing questions may not be related to each other, the rating of one question does not affect other questions³. A question must achieve 100% inter-rater reliability of covering the standard, or it will not be used. Questions which meet this criteria will be “standards-valid graphing questions”, and can be used during the task analysis.

5.2.2 Results

In practice, it was much simpler to be inspired by graphing questions from several sources, and create our own graphing questions for validation. First, since the questions are generated by the research group, there is no ambiguity of copyright. Second, many questions do not directly align with a particular standard, or are too complex, with multiple sections. After some practice, it was easy to create a to-do list of questions, such as “two reading questions for GS 5,” and produce those questions in about 20 minutes.

Questions were created in Microsoft word, in a format that would both be accessible to most students with low vision, and convertible to a braille format for blind students. In this format, the questions can be easily used for later evaluation (phase 2, study 5, and phase 3). A sample of questions used for this study, along with a braille example, are available in Appendix B.

5.2.3 Discussion

This process resulted in a list of graphing questions based on the standards identified in study 1. The goal was to cover reading and writing questions for all of the standards, with more than one question per type. Therefore, $18 \times 2 \times 2 = 72$ graphing questions were required. Most of these have been completed and evaluated with the task analysis portion of this phase.

³In practice, the questions will be very similar, such as “Graph (4,2)” and “Graph (-1,3)”, but, other than the graphing standard, there is no categorization of specific questions, so there is no way to capture how these questions are related.

It is fair to state that answering these graphing questions correctly demonstrates that a student understands a graphing standard. However, there are an endless number of questions that could be created for any standard. Therefore, many categories of questions will not be represented in the examples. However, the graphing questions will be sufficient to represent every standard, for both reading and writing.

Selecting the right sort of questions was difficult. In pilot testing of the questions with students, we found many questions used unfamiliar words, or asked for several answers. After some practice, it became easier to develop appropriate questions. Therefore, this study benefited from an iterative approach.

5.3 Identifying the Steps Needed to Answer the Graphing Questions

Now equipped with an army of graphing questions covering all graphing standards, it is time to find out how students should answer these questions.

5.3.1 Study Design

A task analysis method called task analysis for problem solving (TAPS) [?] was used to find the steps required to complete a question. TAPS is an iterative, turn-taking process for defining a task. In the first iteration, a subject matter expert solves problems, while the researcher takes notes. During this session or a follow up, the researcher asks probing questions about the specific steps and motivation behind the actions taken. At the final session, the *researcher* solves similar problems, and the subject matter experts critique the researcher's solution in terms of alignment with the earlier solutions. In the end, a structured definition of the task is generated out of the 2-3 sessions.

Experts for this study were teachers who know how to teach mathematics and have experience with visually impaired students. On the first day, 2-3 experts were given a graphing problem, and asked to solve it. They were encouraged to use a think-aloud protocol, speaking to a researcher. At the end of solving the question, the researcher cleared any open questions about what steps the experts did, or why they did it. A different question (possibly related to a different standard) was shown, until time ran out in the one-hour session. After the first session the researcher created a list of steps for each question, specific

Table 23: The component behaviors that need to be supported by the system, in order to allow students to answer all of the graphing questions. “Estimate Values” means estimating the current position, between two tick marks with known values. “Filled Region” means setting a region larger than a point that represents the values; this is for number lines only, such as a number line of $x > 3$.

| Object | Dimension | Action | | | |
|-------------------|-------------|--------|-----|------|--------|
| | | Find | Add | Edit | Remove |
| Origin | Graph | X | | | |
| Tick Mark | Both | X | | | |
| Tick Mark Label | Both | X | | | |
| Estimate Values | Both | X | | | |
| Point | Both | X | X | | X |
| Label | Both | X | X | X | X |
| Point Open/Closed | Number Line | X | | X | |
| Filled Region | Number Line | X | X | X | X |

enough to solve that question, but also general enough to solve similar questions.

During a follow-up session, the researcher solved a similar question, with the teachers observing and pointing out any changes. This was repeated in future sessions until there were no more questions. In 7 TAPS initial and follow-up sessions, all graphing questions were solved. Based on the number of questions and a one-hour session length, each question took about five minutes to solve.

The level of detail of this task analysis was one level above the perceptual activities. For example, for the question “Graph (2,3)”, teachers may suggest to start by finding the origin, then move along the tick marks to the right (since 2 is positive) until at the tick mark for number two. For this process to work, the graph must enable users to somehow perceive the origin, and be able to move to the right, and detect where their current location is. The focus is on the conceptual tasks, instead of the perceptual tasks (e.g. looking at the graph).

5.3.2 Results

A comprehensive list of the standards, steps, and component parts are available in Appendix C. Table 23 has a list of the component behaviors that must be supported in order to allow students to answer all of the graphing questions. Table 24.

Table 24: The behavioral components' relationship to the graphing standards.

| | | Action | | | | | | | | | | | | | | | | |
|-------------------|----|-------------|----------------|----------------------|-----------------|------------|-----------|--------------|------------------|-----------------|------------|-----------|------------|--------------|--------------------|-------------------|--------------------|----------------------|
| | | Find Origin | Find Tick Mark | Find Tick Mark Label | Estimate Values | Find Point | Add Point | Delete Point | Find Open/Closed | Set Open/Closed | Find Label | Add Label | Edit Label | Delete Label | Find Filled Region | Add Filled Region | Edit Filled Region | Delete Filled Region |
| Graphing Standard | 1 | X | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 2 | | X | X | X | X | X | X | | | X | X | X | X | | | | |
| | 3 | | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 4 | X | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 5 | | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 6 | X | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 7 | X | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 8 | | X | X | X | X | X | X | | | X | X | X | X | | | | |
| | 9 | X | X | X | X | X | X | X | | | X | X | X | X | | | | |
| | 10 | | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 11 | X | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 12 | | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 13 | X | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 14 | | X | X | | X | X | X | X | X | X | X | X | X | X | X | X | X |
| | 15 | X | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 16 | X | X | X | | X | X | X | | | X | X | X | X | | | | |
| | 17 | X | X | X | | X | X | X | | | X | X | X | X | | | | |
| Count | | 10 | 17 | 17 | 3 | 17 | 17 | 17 | 1 | 1 | 17 | 17 | 17 | 17 | 1 | 1 | 1 | 1 |

5.3.3 Discussion

There are two important outcomes of the task analysis. For each question, there is an ordered list of steps to complete. Taken together, the questions represent the graphing standards. Therefore, their steps represent the pieces of the entire graphing activity necessary for demonstrating competence in graphing standards. This information can be used to build a system that enables students to complete all of the pieces of the graphing activity.

It is remarkable how few steps there are. Over 17 standards, reading and writing, there are only 17 steps. Furthermore, only 10 steps are used in all standards, and up to 11 steps are used in the 16 standards other than GS.14 (number line inequalities). A critical part to many of the steps is point estimation. Based on the findings of phase 1, along with speech and a simple menu interface, these steps should be straightforward to develop.

A key component of later development is ensuring that the steps that the students can do with the new tool are very similar to the steps that the students can do with the

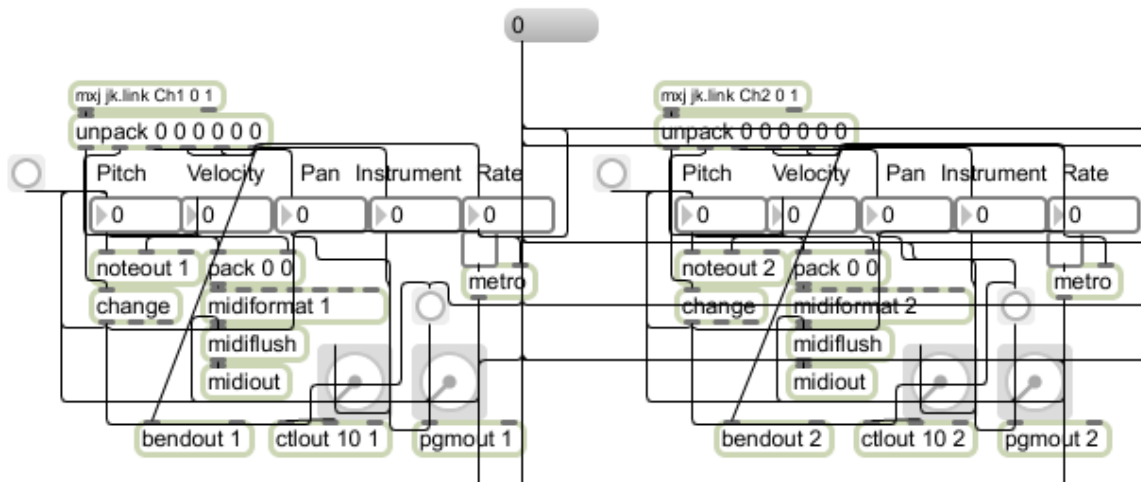


Figure 40: The Max/MSP receiver. Each group is a single channel. Two channels are shown, 16 are available (one for each MIDI track), but only a maximum of four were used.

current tool. Therefore, during this process and the summary reporting, it is important to adequately characterize the steps.

At this point, the process is two conversions away from the original standards. The standards have been represented by questions, and these graphing questions have been represented by their component steps. In a sense, the standards can be characterized by the steps of the graphing questions. In other words, if a tool enables the steps necessary for a graphing question based on a standard, the student is being equipped with a tool that lets him demonstrate competency in the standard. However, a more comprehensive validation will be left to the last phase, in evaluation.

5.4 *Reconstruct the Steps in an Auditory Graphing Technology*

Software development is a complex method. It involves coding skills, personnel management, debugging, and stakeholder meetings. Based on 10 years experience programming in industry and research, I orchestrated the development of the Auditory Graph Model (AGM), Accessible Graphing Engine (AGE), and Graph and Number line Interaction and Exploration system (GNIE). A description of these technologies is covered in more detail in chapter 3.

Due to the dependency on existing Java systems, and the portability of Java, the system

core is in Java. During the phase 1 psychophysical studies (Chapter 4), it was sufficient to use Java for the audio processing. In those four studies, only one sound representation was running at a time. Later evaluations used two and three sounds in production, and up to four in development (specifically x-axis, y-axis, and either radial point distance or x and y distances each having a sound simultaneous with the others). While the latency remained low on many systems, the load was high enough to crash the system or the Java Virtual Machine.

Since Java could not handle the audio processing, Max/MSP was used as an alternative. This program/language has a reputation for low-latency audio processing. In addition, using maxlink, it was possible to send messages to Max/MSP from Java. The Max/MSP program, ProcPlayer, played sounds based on messages with information about channel, instrument, pan, pitch, pitch bend, rate, and volume. See Figure 40.

During the long pilot study of the evaluations (phase 3, discussed in Chapter 6), several iterations of the system were developed. Student feedback from a participatory design process made the system more functional, usable, and accessible.

5.5 General Discussion

The SQUARE method is a systematic, evidence-based method for generating a relevant tool out of education standards. During standards identification, the high-level requirements were outlined. In question selection, exemplars of the requirements were catalogued. During the Answer portion, a high level TAPS task analysis was conducted to obtain the steps. After the system was developed in Reconstruction, it was evaluated in a side-by-side comparison of contemporary tools. As applied to the sixth grade CCS for Mathematics, it provides a: list of graphing standards, example graphing questions, steps needed for sixth grade graphing, a tool for graphing in sixth grade, and proof that the tool is suitable for the questions, standards, and the grade.

SQUARE could be applied to other grades or other topics. The steps are of particular importance, since they outline the design requirements for new systems. This can inform the development of any technology for the grade. It could also be used to see the development

of graph literacy (and other topics) over the K-12 standards and curriculum.

One limitation of this approach is question comprehensiveness. While every standard will be evaluated, there are endless of possibilities for graphing questions. Clearly not every graphing question can be evaluated. Therefore, not all of the steps possible for solving the standard will be uncovered. The scope of this dissertation will be to identify one type of reading question and one type of writing question for each standard. While other graphing questions (and steps) are possible, a student could learn and demonstrate competency in the standard by using the limited types of graphing questions presented.

The learning scientist may disagree with the approach used in phase 2. The approach use has a behaviorist slant [?], including the use of task analysis, observable behaviors, and categorical selections. Cognitive and situational approaches can inform observers about learning further (as explained in [?], and will be explored further in the more ecological evaluation withs students in phase 3. This work is not designed to change the content of curriculum. Instead, it is designed to provide alternative presentation formats to the same content, so that more students can participate.

5.6 Conclusion

GNIE is based on graphing standards and has validity within SQUARE. The final phase evaluates the system with students and teachers.

CHAPTER VI

PHASE 3: EVALUATION IN CLASSROOM AND TESTING SCENARIOS

System evaluation occurred in two major parts: evaluating performance of students using Graph and Number line Interaction and Exploration system (GNIE) in classroom settings and in testing situations. Each extend Phases 1 and 2 to provide a more ecological assessment of GNIE. Evaluation was completed in Spring and Summer 2012.

6.1 Classroom Evaluation

Graphing tools for K-12 mathematics will mostly be used in classrooms and activities like homework. Therefore, it is critical to evaluate whether GNIE could be used in classroom situations. A short-term comparison of contemporary graphs alongside GNIE would enable informative observation and engage discussion of the tradeoffs between the different graphing formats.

A classroom evaluation opens the door for comparisons on a social level, one of three major perspectives on learning [?]. New questions about the use of the tools and its impact on teacher-student and student-student interaction can be observed and discussed.

There are the following questions for this study:

Q_1 How can the basic concepts of auditory graphs be learned quickly?

Q_2 What is the impact of an auditory graphing tool on solving graphing problems in class?

Q_3 What are student and teacher opinions of the auditory graphing tool?

6.1.1 Study Design

These three questions were answered through a four-day workshop at a summer camp for visually impaired students in July 2012. Q_1 was primarily addressed through evaluation

of the Navy game. Q_2 will be discussed in terms of findings from the pilot studies and observations during the workshop. Q_3 is reported through an exit survey from the workshop.

6.1.1.1 Apparatus

The classroom evaluation will be a recreation of the situation found at schools for the blind, with small class sizes and many resources. This gave an opportunity to evaluate the system in an ideal classroom situation.

6.1.1.2 Participants

30¹ visually impaired students attending the Center for the Visually Impaired (CVI) STARS summer program participated in this study. Students were in middle and high school, so they were familiar with sixth grade mathematics content².

Human subjects research approval for a participating student occurred on three levels. First, both CVI and the Georgia Tech Institutional Review Board have approved the research³. Second, parents were informed of the study in a packet about the CVI summer program, and were asked to sign consent. Third, students were asked to provide a verbal assent at the start of the first session.

6.1.1.3 Task

The CVI summer program has blocks of activities throughout the day. Student-participants came in two groups, with each session lasting about 75 minutes. There were four sessions, and most students attended every session. In order to further break down the group, the students were further divided in half, and given a "team" color, either blue, green, purple, or red. Students also used a particular computer, which was named after one of the four classical elements or a season of the year. Students were then tracked based on this combination of color and computer name, such as "Blue Fire" and "Purple Spring".

The CVI summer program has blocks of activities. For most blocks, there are three

¹verify count

²Although the Common Core Standards (CCS) and Common Core Georgia Performance Standards (CCGPS) are not used in Georgia until 2014, there is a great deal of overlap, so most 7-12 grade students should be comfortable answering questions based on the sixth grade CCS.

³Georgia Academy for the Blind (GAB) and CVI provided approval during the pilot studies.

activities. Students are divided into three groups, and rotated between the activities after one hour, for three hours. With 25 students, the summer program will have about 8 students per group. Therefore, this study will have about 8 students in each session, with 3 groups of students.

6.1.1.4 General Task

The study is presented to students in four sessions, each one hour long. During the first session, students began by providing assent. An introduction to graphs was then provided, using contemporary tools such as high-contrast graphs with large font or the Graphic Aid for Mathematics. Near the end of the session, students practiced finding a point alone and on a number line with an adaptation of the AudioFitts program used in Phase 1 (Section 4.1.1).

During the second day, students were provided with a formal introduction to the GNIE program. They practiced each of the steps described in the Answers study for Phase 2 (Section 5.3 and Appendix C), and then answered example graphing questions extracted from the Phase 2 Questions study (Section 5.2 and Appendix B).

The third and fourth days both had graphing and discussion activities, counterbalanced between groups. On each day, the students used a technology to answer graphing questions. These questions were based on a lesson designed by a mathematics teacher familiar with teaching visually impaired students. If possible, the teacher also taught the content during the session. The teaching component lasted about 45 minutes. For the remaining 10-15 minutes, students gave feedback about the session and technology, and fill out a short survey with Likert and open-ended questions about the technology. The half of the group used paper on session 3 and GNIE on session 4, while the other half used the tools in the reverse order. See Table ??.

The purpose of counterbalancing is primarily to evaluate the use of the tools with different types of graphs. Since there are only two lessons, one will focus on standards and questions about number lines, and the other will focus on graphs. It would be more difficult to compare paper number lines with 2D auditory graphs, in terms of observation and student

feedback. Instead, half of the students experience one combination and half experience the other.

Data collection occurred in three ways. First, the sessions were recorded, in notes, audio recording, video recording, pictures, and materials from students. The audio recordings were transcribed. Then, a story of the sessions was reconstructed. It included quantitative measures such as the number of problems completed, and qualitative data such as student comments on using a graphing tool. The goal of the data analysis was to paint a picture of what happened. Much of the analysis is loosely defined, as there is no core theory to be uncovered or specific hypotheses for the question on classroom impact.

Second, the students also completed a survey. It included their level of impairment, graph experience, computer experience, and feedback about the sessions and GNIE. As described in Section 2.4.1.1, student opinions are an important part of their motivation to learn the materials; their opinions will be established with this survey. Their subjective ratings of performance with GNIE and the paper tools were compared with the quantitative results.

Third, the students practiced auditory graphs with the Navy game. Over a period of five minutes, they tried to select as many ships as they could find. Their scores were logged and compared to see if there was an improvement in scores.

6.1.2 Pilot Study

A pilot study was conducted in spring 2012 with middle and high school students at after school programs GAB and CVI. The workshop lasted 11 weeks, and attempted several types of evaluation, including testing evaluation discussed in Section ???. This longer process resulted in an improved GNIE system from productive participatory design. However, the project was loosely designed, so follow-up is necessary for cleaner results.

One of the potential challenges was maintaining an interest in solving mathematics problems over the course of several weeks. In order to provide motivation, and to collect information on training, a game was invented to encourage point estimation. The game, called “Navy”, was like the old peg game “Battleship” in the sense that ships were at

particular coordinates. However, in Navy, there is one dimension (a horizontal number line), and the ships are moving. Participants must move left and right to find ships (graph points), identify their location (tick mark label), and indicate on the 0-9 range the position of the ship. Then an explosion sound was played and the ships near the target label disappeared. The auditory and visual displays are similar to the AudioFitts displays used in Phase 1, except with a nautical theme.

Low vision students could look at the graph, but were sometimes asked to obscure their graph so they had to listen to the sounds for the game. All students appeared to be highly engaged with this activity, and many scores were in the hundreds of points within 10-20 minutes⁴. See Figure 27 for a graphic of what the Navy game looked like.

During the last two sessions at GAB, a teacher presented an hour-long lesson. This lesson was based on one or more of the validated graphing standards from Section 5.1, with one session focusing on number lines and the other on graphs. The teacher presented the same content to two different groups of visually impaired students. One group used current tools that the teacher provides. The other group used GNIE, and any supplementary tools the teacher suggests⁵. On a separate day, a different teacher presented other graphing content to the same groups. Except this time, the two student groups used the other technology⁶. The general idea is discussed in the proposed study design and shown in Days 3 and 4 of Table ??; for the pilot study, there were only two groups, and on the second day the groups were combined, but the technologies used remained counterbalanced.

6.1.2.1 Pilot Results

Several results are clear. Students were able to complete the lessons with both paper and computer-based tools. In general, students could use GNIE to answer graphing questions, and rated it similarly but slightly lower than existing tools. In terms of teacher preparation, the ease of plotting points and mass producing graphs with GNIE appears to provide a

⁴Scoring was +2 for each ship hit, -1 for a miss, and -1 for letting a ship get across the screen.

⁵While it is tempting to put GNIE separate from the other tools, it is not realistic, and puts the new technology in a position to compete against *all* of the existing technologies. Instead, the teachers used GNIE as the primary tool, and had other tools available.

⁶Due to a scheduling issue, students on the second day were combined into one group. Each student used the technology-set they hadn't used on the first day.



Figure 41: Students training on point estimation on a number line.

distinct advantage to preparing tactile graphics.

One striking difference is the level of communication when using the different tools. On the first session with the first group, students talked a lot, both with the teacher and with each other. Topics were about the lesson or other topics. With the second group, who used GNIE, students did not talk with each other. During the second session, the talkative students were using GNIE, and stopped talking, except for a couple of questions with the teacher. The students in the second session who used paper talked with each other and the teacher more often. Students using GNIE reported hearing others, but not interested in talking.

Three potential causes of the silence are: use of audio, use of headphones, and use of computers. In order to evaluate the cause of the talking differences, we asked two GNIE-using students in the second session to take off their headphones, but keep audio on, for

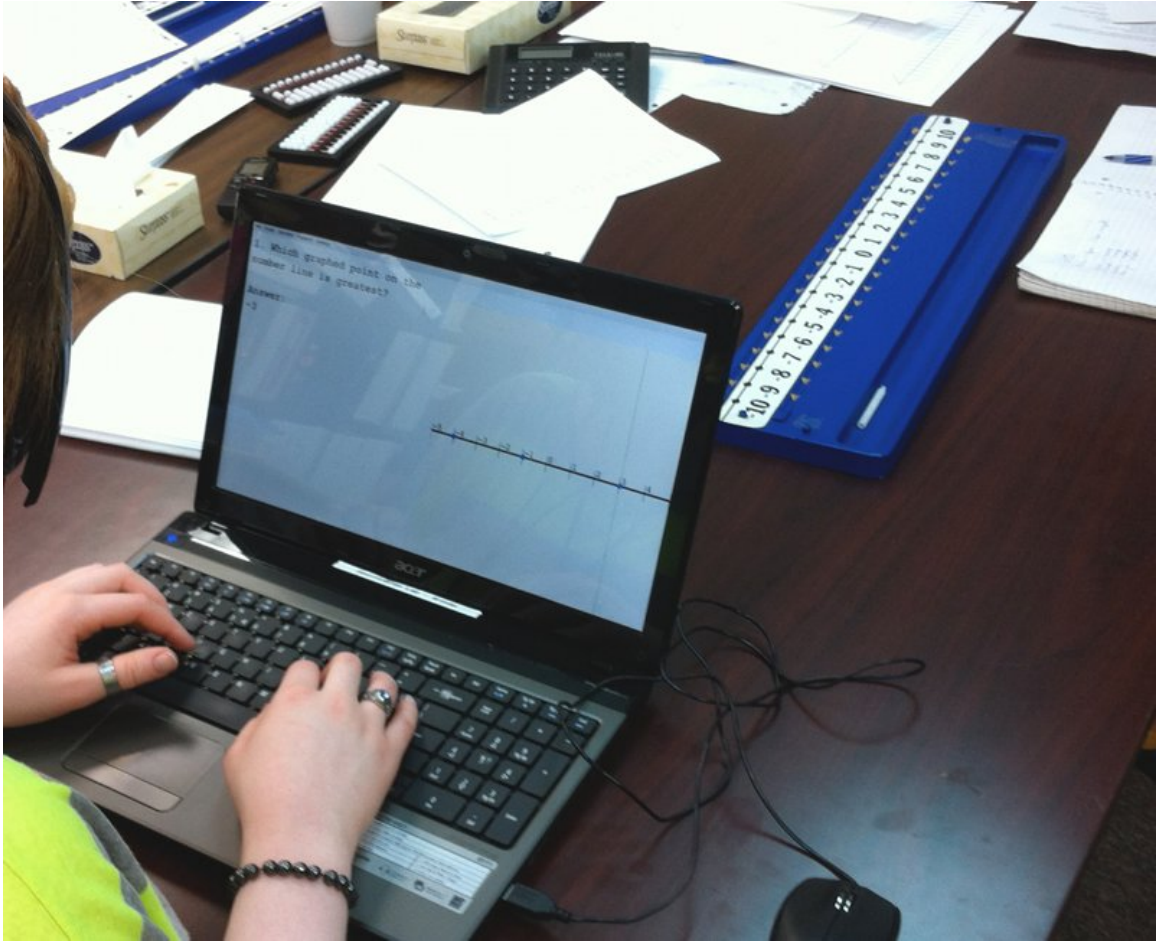


Figure 42: A student using GNIE to solve a graphing problem.

about 10 minutes. We also asked them to turn audio off and keep their headphones off for another 10 minutes. In both cases, talking remained low with the students. It appears that using the computer strongly affects communication. It is unclear whether this affects learning, however. The students appeared to solve problems more independently with the computer, and the paper-using students worked together more; but each group also completed the graphing problems in about the same amount of time. From a behavioral and cognitive perspective (e.g. [?]), not much may have changed, but from a social perspective, adding computers (or perhaps headphones or audio) has a big impact on the student.

The classroom pilot shows promising results, in terms of teacher preparation, student achievement, and stakeholder opinions. An interesting finding related to the communication

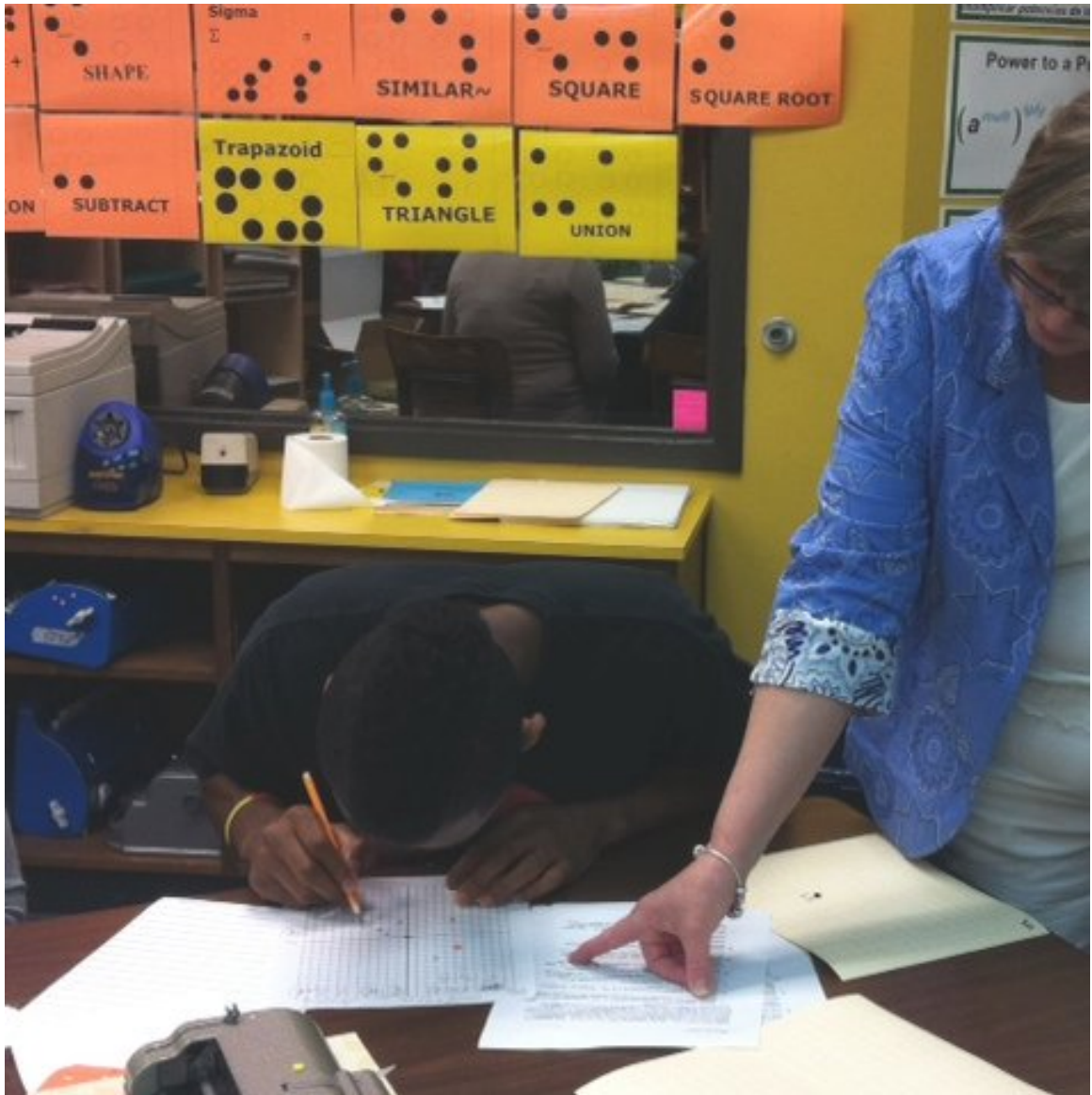


Figure 43: A teacher working with a student on paper graphs.

between students and teachers based on the tools used raises questions about negative impacts of auditory graphs, but it appears this impact is computer-based and had no measurable effect on student performance.

While GNIE and auditory graphs are the featured technology, it is critical to include existing technologies in the evaluation. Students and teachers are familiar with these technologies and no single tool will replace them. An emphasis on the new technology, while providing the old technologies, worked well for the classroom comparison study. Students would use the other tools for insight, but often sparingly.

6.1.2.2 Navy Game Task

The Navy Game study was conducted with one day of training and two days of data collection. During the first class session (day 1), students learned the rules and practiced with the game for about 10 minutes. On day 2, and either day 3 or 4 (depending on their group), students played the Navy game for a minimum of five minutes. Participants were allowed to look at the screen, use headphones, hold a mouse, and press the number keys to target ships (for more details on the Navy Game, see Section ??). Participants were motivated by the game itself, and that their color team would win a prize if they had the highest total score for the day. The top player would also get a prize. Prizes were additional snacks at the end of each day.

The Navy Game is a training tool for the basics of number line (and subsequently graph) movement. We wanted to know: how will students improve with the Navy Game? Fast improvements, within five or ten minutes of training, would indicate that low vision and blind students can quickly learn the basic actions for navigating the graph. Of course, playing a game is not the same as a real graph, so the results would not be directly comparable to graphing, but the learning effect could be measured.

As described earlier, the Navy Game increases in difficulty every minute (there are more ships). Therefore, the potential range of score changes each minute, and minutes are not directly comparable by score. However, averages over a set of minutes, or particular minutes from the start, are comparable to these chunks in other game sessions. Therefore, we

examined performance between sessions for each minute, and average improvement between days. We hypothesized that:

H_1 For every minute, there would be an improvement in scores from the first day to the second day.

H_2 Between days, there will be a significant improvement in scores.

The familywise alpha was set to 0.0083, using the conservative Bonferroni correction: $\alpha = 0.05/6 \approx 0.0083$.

6.1.3 Results

6.1.3.1 Navy Game

During the first minute, there was a descriptive improvement in scores from -0.42 to 3.33, but no significant difference, $t(11) = 2.181$, $p > 0.0083$. During the second minute, there was a significant increase in scores, from 2.08 to 15.42, $t(11) = 3.406$, $p < 0.0083$, $d = 1.97$, or 13.34 points. During the third minute, there was a significant increase in scores between days, from 13.75 to 30.33, $t(11) = 4.593$, $p < 0.0083$, $d = 2.65$, or 16.58 points. During the fourth minute, there was a significant increase in scores between days, from 35.33 to 60.42, $t(11) = 3.663$, $p < 0.0083$, $d = 2.11$, or 25.09 points. During the fifth minute, there was a significant increase in scores between days, from 63.42 to 100.33, $t(11) = 4.121$, $p < 0.0083$, $d = 2.38$, or 36.91 points. Between days, there was a significant increase in overall score, from 114.17 to 209.83, $t(11) = 4.252$, $p < 0.0083$, $d = 2.45$, or 95.66 points. Table 25 shows the differences in Navy Game performance.

It is not particularly satisfying to treat each minute independently. While comparisons are problematic due to the changing level of difficulty and number of ships in the game, a general picture of improved performance does take shape when looking at the percent of total scores participants had for each minute. Figure 44 shows this relationship.

6.1.3.2 Classroom

Low vision and blind students both easily used GNIE, along side the tactile and visual graphs. Partners worked together, and each session was alive with talking and problem

Table 25: Results from the Navy Game experiment. "D" stands for "day", "R" stands for "round", and "Cum" stands for "Cumulative" (the sum of all rounds for the day). There were significant improvements in 5 out of six of the tests, showing clear signs of students improving on navigating toward targets.

| y_1 | y_2 | \bar{y}_1 | s_{y_1} | \bar{y}_2 | s_{y_2} | $t(11) = ?$ | sig? | d | $y_2 - y_1$ |
|-------|-------|-------------|-----------|-------------|-----------|-------------|------|------|-------------|
| D1R1 | D2R1 | -0.42 | 7.24 | 3.33 | 3.34 | 2.181 | no | 1.26 | 3.75 |
| D1R2 | D2R2 | 2.08 | 18.19 | 15.42 | 6.89 | 3.406 | yes | 1.97 | 13.34 |
| D1R3 | D2R3 | 13.75 | 23.43 | 30.33 | 15.23 | 4.593 | yes | 2.65 | 16.58 |
| D1R4 | D2R4 | 35.33 | 34.94 | 60.42 | 16.70 | 3.663 | yes | 2.11 | 25.09 |
| D1R5 | D2R5 | 63.42 | 45.44 | 100.33 | 23.76 | 4.121 | yes | 2.38 | 36.91 |
| D1Cum | D2Cum | 114.17 | 124.09 | 209.83 | 61.73 | 4.252 | yes | 2.45 | 95.66 |

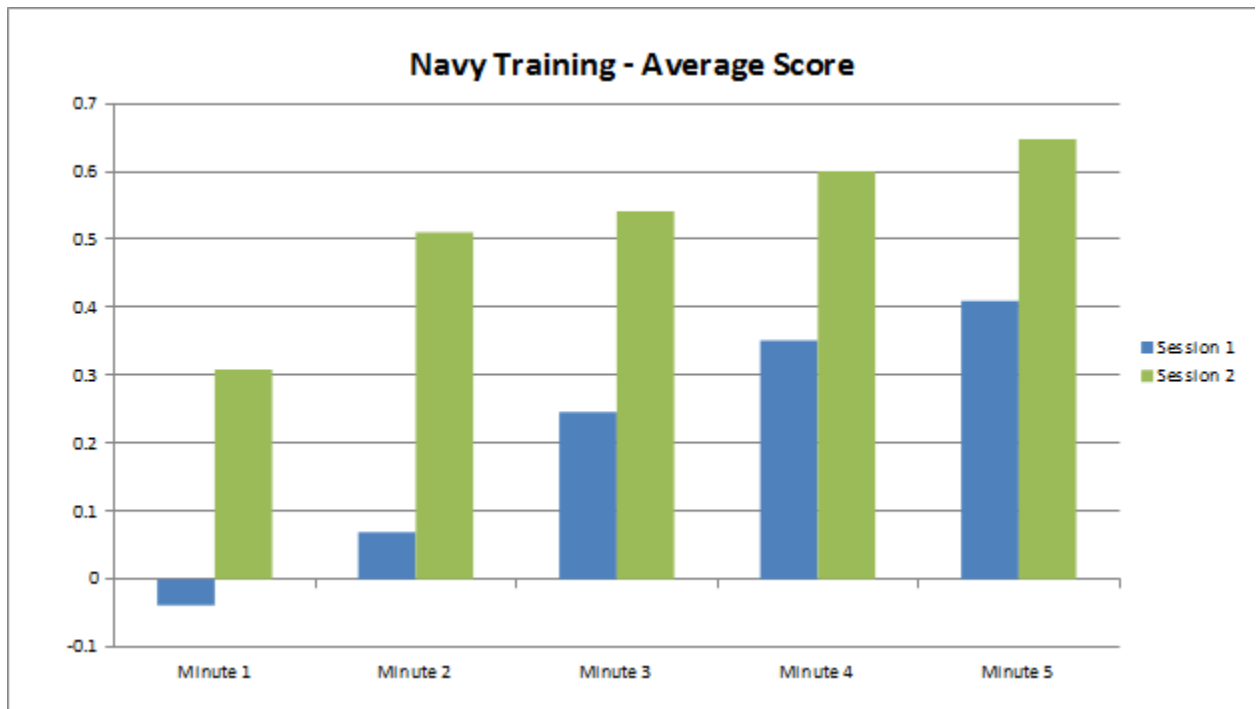


Figure 44: Mean scores of the Navy Game as a percent of the total possible score for each round. Note that the scores can be below zero.

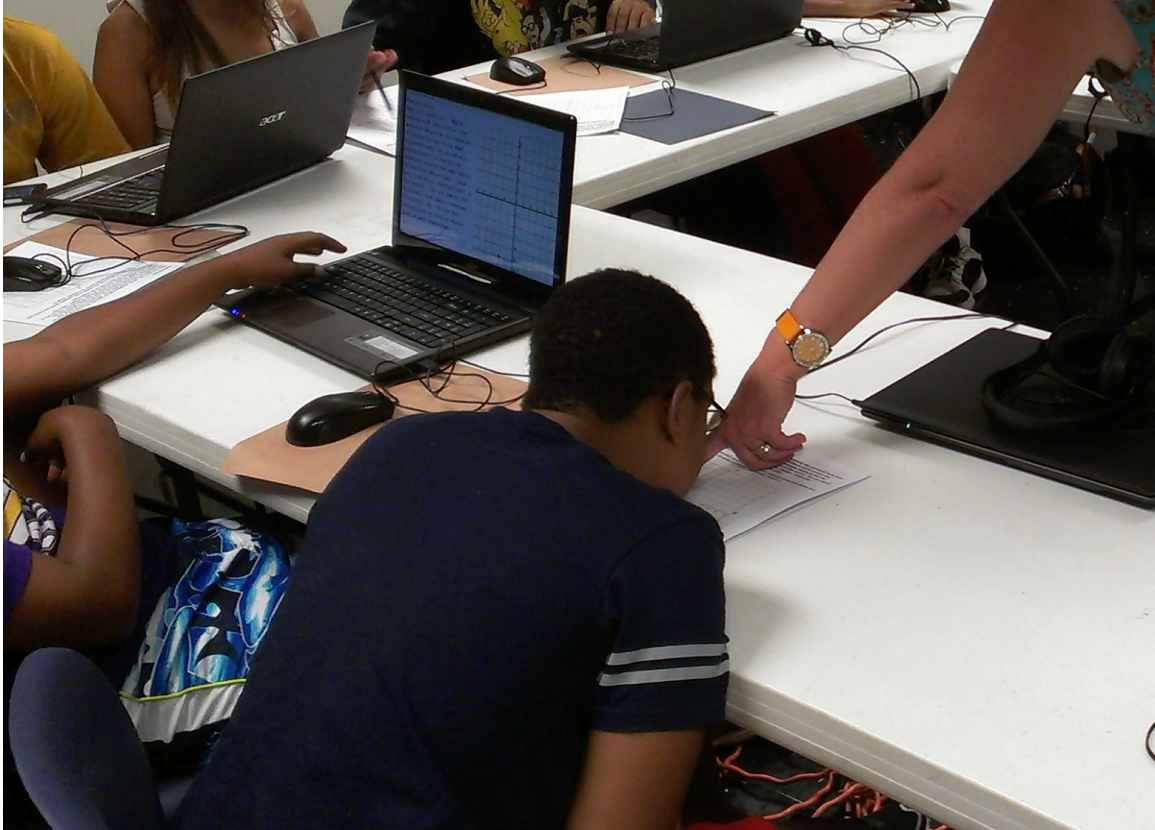


Figure 45: Students worked in pairs of two, one of which used auditory graphs on the computer and one who used visual or tactile graphics. The teacher talked with the students, too.

solving.

Surprisingly, the differences in classroom interaction seen in the pilot did not manifest themselves in this study. It may be that the situation in the pilot was too controlled. Since everyone was using GNIE, there was a behavioral change. When people using the computer were working with people without a computer, this phenomenon did not occur.

6.1.3.3 Student Feedback

Table 26 shows the results of the exit survey from the classroom activity. Participants were generally confident in their capabilities to solve the graphing problems with either tool. However, participants may have higher confidence in their current tools for solving graphing problems in classroom situations.



Figure 46: Low vision students could magnify GNIE.

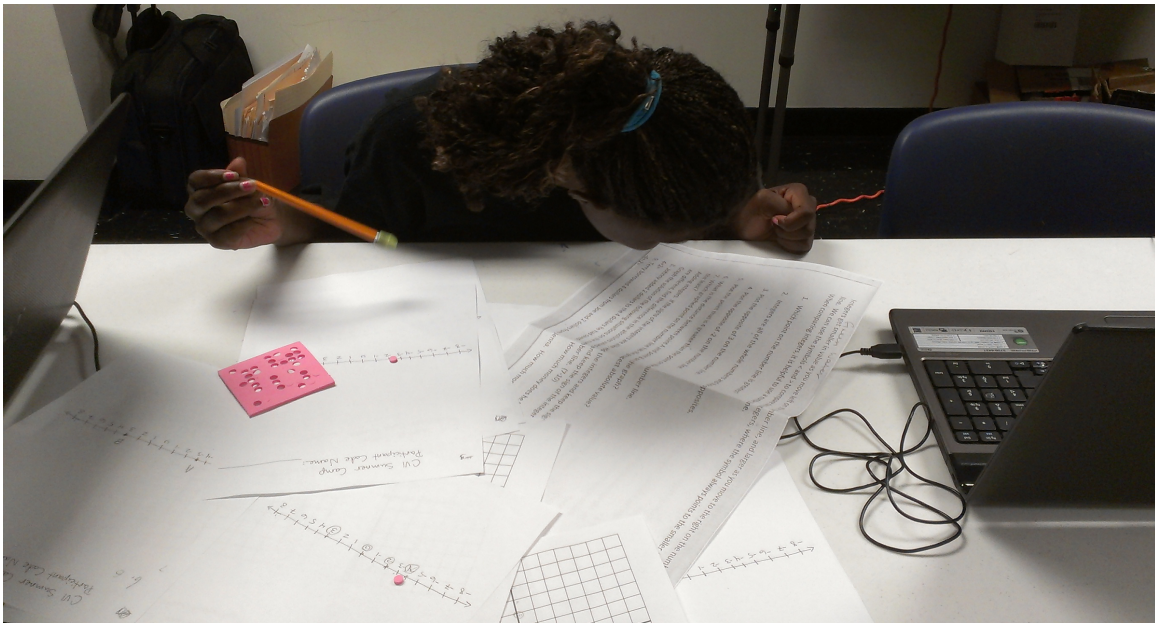


Figure 47: This student is using a combination of visual graphs and tactile graphics to answer the questions.

Table 26: Scores from the exit survey during the summer camp (n=19). The survey asked participants to "answer with a scale of 1 to 5, where 1 is strongly disagree, 2 is somewhat disagree, 3 is neutral, 4 is agree, and 5 is strongly agree."

| Statement | 1 | 2 | 3 | 4 | 5 | Mean | Median |
|---|---|---|---|---|----|------|--------|
| In general, number line questions are difficult. | 7 | 5 | 3 | 2 | 2 | 2.32 | 2 |
| In general, graph questions are difficult. | 6 | 7 | 4 | 1 | 1 | 2.16 | 2 |
| I understood the graphing problems with paper/-tactile graphs. | 2 | 0 | 4 | 5 | 8 | 3.89 | 4 |
| I understood the graphing problems with computer graphs. | 4 | 1 | 4 | 2 | 8 | 3.47 | 4 |
| I needed the sound to use the computer graphs. | 5 | 2 | 3 | 2 | 7 | 3.21 | 3 |
| I needed my vision to use the computer graphs. | 5 | 4 | 2 | 2 | 6 | 3.00 | 3 |
| I needed my vision to use the paper/tactile graphs. | 3 | 5 | 0 | 3 | 8 | 3.42 | 4 |
| I preferred to be able to hear the sounds with the computer graphs. | 4 | 2 | 4 | 2 | 7 | 3.32 | 3 |
| Between the paper graphs and the computer graphs, I think I was faster on paper. | 3 | 3 | 0 | 7 | 6 | 3.53 | 4 |
| Between the paper/tactile graphs and the computer graphs, I think I had more graph questions correct on paper. | 2 | 0 | 5 | 5 | 7 | 3.79 | 4 |
| Between the paper graphs and the computer graphs, I think I was faster on the computer. | 5 | 8 | 1 | 1 | 4 | 2.53 | 2 |
| Between the paper/tactile graphs and the computer graphs, I think I had more graph questions correct on the computer. | 4 | 8 | 2 | 2 | 3 | 2.58 | 2 |
| I usually had my headphones off. | 8 | 3 | 3 | 1 | 4 | 2.47 | 2 |
| I could often use paper/tactile graphs in my regular math class. | 1 | 0 | 1 | 4 | 13 | 4.47 | 5 |
| I could often use computer graphs in my regular math class. | 8 | 4 | 1 | 5 | 1 | 2.32 | 2 |
| It was easy to understand the meaning of the sounds on the computer graphs. | 4 | 2 | 1 | 7 | 5 | 3.37 | 4 |
| I had to concentrate on the sounds on the computer graphs. | 2 | 1 | 2 | 6 | 8 | 3.89 | 4 |
| Paper/tactile graphs were helpful in solving problems. | 2 | 4 | 3 | 5 | 5 | 3.37 | 4 |
| Computer graphs were helpful in solving problems. | 3 | 5 | 4 | 2 | 5 | 3.05 | 3 |
| Paper/tactile graphs were fun to use. | 4 | 3 | 7 | 2 | 3 | 2.84 | 3 |
| Computer graphs were fun to use. | 5 | 0 | 7 | 3 | 4 | 3.05 | 3 |
| Paper/tactile graphs were annoying to use. | 5 | 4 | 7 | 1 | 2 | 2.53 | 3 |
| Computer graphs were annoying to use. | 6 | 3 | 3 | 3 | 4 | 2.79 | 3 |

6.1.4 Discussion

In the Navy Game, except for the first minute, there were large, significant differences between the days. It appears that students quickly improved their score, and that the basics of navigating a graph like that in GNIE can be learned in under half an hour. Recall that the ships were moving in the game, and there were many points; a regular 6th grade graphing problem would be much simpler.

Students in the classroom showed that they could work together to solve graphing problems. Neither tactile, visual, or auditory graphs blocked answering the questions. The surveys showed that students had mixed opinions of the technology, perhaps in part since it had been a short introduction, and the tool is not ready for classroom use in terms of training materials, technology debugging and polish, and curriculum integration. Nevertheless, they appear to have been able to use it in the classroom setting.

6.2 *Examination Performance*

One challenge in education is providing tools that can be used in both classroom and testing situations. In the classroom, a student can be guided by the teacher and other students in the use of a technology, and also monitored for doing his share of the work. In testing situations, strict technology guidelines ensure fair comparisons between scores of students. When providing testing accommodations, alternative tools must provide access but also not enable the student to get an advantage over his sighted peers [?].

In this dissertation, identifying the specific construct being evaluated on a test question is based on conducting the same conceptual steps as others. As this was completed as part of the SQUARE task analysis in Section 5.3, GNIE should enable users to answer questions in the same way. However, it may be the case that alternative, easier ways of solving the problem are also available.

One way to evaluate the ease of a situation is to test whether a disabled student improves, and whether a sighted student improves. The disabled (blind) student should *perform higher or equally well* with the technology when compared with an alternative; otherwise the new system is not a better accommodation than the existing approach. The non-disabled

(sighted) student should *perform equally well or lower*; otherwise the new system provides an advantage over the existing (standard) approach. In order to compare populations that are more similar, low vision participants were used as comparisons to blind participants.

This section compares examination scores between visual, tactile, and auditory graphs. It shows how surprisingly similar the different formats are, and the discusses implications for converting between formats on tests. The first part introduces the evaluation method, including the novel "over-under" evaluation. The second part describes the results and evaluates the test scores in terms of over-under. The third part adds the perspective of test difficulty, showing how most questions, when more difficult in one format, are also difficult in the other format.

6.2.1 Study Design

Allowing computerized forms of testing increases the diversity of testing materials for blind students, when the scores stay the same or improve, but not if the scores get lower. Therefore, the comparison should ensure that scores are not going below that of the current technologies. In addition, sighted students should not outperform their scores when using visual tests, while blind students may over-perform. Significant values are important, but more important is the observed effect size. While some effect size can be attributed to learning a new system, if the effect size is large in the wrong direction, then the system may not be suitable for students taking a test.

Examination results focused on student accuracy on the graph questions (correct and incorrect answers). All questions were be multiple choice, "graph reading" questions (see Section 5.2). Writing and open-ended reading questions are an important part of graph literacy, but interpreting the answers can be subjective and may lead to more bias. In addition, many formal mathematics tests only evaluate students using multiple choice, probably due to the same concerns about subjectivity and bias.

There were the following questions for the examination scores:

Q_1 What is the effect of graph format on the graphing test performance of tactile graph users?

Q_2 What is the effect of graph format on the graphing test performance of visual graph users?

Q_3 What is the relationship between test scores on particular problems for visual graphs and auditory graphs?

Q_4 What is the relationship between test scores on particular problems for tactile graphics and auditory graphs?

The first two questions compare the mean differences of auditory and either visual or tactile graphics. Ultimately, the goal is to provide an appropriate testing accommodation, so, as discussed earlier, the scores must meet two criteria. First, the auditory graph scores must be approximately equal to the tactile graph score, so that braille users may be able to use either format on a test without a large negative effect. The auditory graph score can also be significantly larger than the braille format, but not smaller⁷. This first test will demonstrate whether blind students can achieve comparable or better scores than before. However, a second test is needed to determine whether auditory graphs provide *too much* support.

Second, the auditory graph scores must be approximately equal to the visual graph score. Low vision and sighted users do not gain an advantage on the test by using the new tool. The auditory graph score can also be significantly lower than the visual format, but not larger. This second test will demonstrate whether blind students are getting an unfair advantage with tools that will use active auditory graphs and similar tools.

Q_3 and Q_4 introduce question difficulty. If students get a particular graphing question correct 80% of the time in a visual format, they should arguably also get it correct about 80% of the time in tactile and auditory formats. If they do not, it suggests that the modality interferes with the problem-solving strategies for that question. If they do, it suggests that solving graph problems is independent of whether the graph is in a visual, tactile, or auditory format.

⁷I make a small accommodation in the test, to account for the low level of experience with auditory graphs.

6.2.1.1 Apparatus

The experiment was conducted mostly in two locations. The first was in a large room in the basement of a local visual impairment work rehabilitation organization, the Center for the Visually Impaired, in Atlanta, Georgia. The second space was a stage at the Georgia Academy for the Blind in Macon, Georgia. In one case, two participants completed the study at a residential home.

The computers were laptops, specifically Acer Aspire 5750Z-4887, with dual-core 2GHz processors and 4GB RAM. The operating system was Windows 7 professional, 64-bit. The aspire laptops were selected in part because they have a full keyboard, including a 10-key and arrow keys. Participants could use the trackpad if they wanted, but often participants used a UBS mouse⁸. Participants wore a variety of headphones, similar to and including Sennheizer HD 202 headphones.

We found it important to add navigational aids to the laptops. Blind and low vision people use a variety of technologies for interacting with a computer. Some are familiar with QWERTY keyboards, some use note takers, and some have little computer experience. In order to facilitate keyboard interaction, important keys were marked on the keyboard. Vincent Martin, a Sonification Lab member with technology rehabilitation training experience, suggested locator dots for a few major keys. As described in the Technology chapter, GNIE uses the 10-key, arrow keys, tab key, backspace key, spacebar, and the letters x and y⁹. We added large dots to the tab and backspace keys, and small dots to the y, down arrow, and middle ("5") key on the 10-key region. This assisted users less familiar with the keyboard while allowing adept users to use the same keyboard.

There are many factors which influence the efficacy of a testing accommodation. On computers, for example, the choice of text-to-speech (TTS) engine for a blind student can have a major impact on their scores. Understanding the synthesized voice and knowing how to navigate with the TTS takes time, and is still being worked out in modern testing.

⁸The trackpad is where people may normally rest their hands, so it was typically disabled for participants who could not see and found it an annoyance.

⁹GNIE allows for inputting regular text in the text box and for graph point labels, but the keys specified have additional meanings.

Our goal, however, is to explore the effect of graph format, not the instructions or answer format.

The test, then, consisted of three parts: an instructions packet, answer sections, and a graph. The instructions and answers format were the same for any particular participant, but varied between participants. For example, low vision participants used visual instructions, and blind participants could use tactile instructions. In addition to auditory graphs, the graph formats presented were either visual or tactile graphs, depending on the participant's level of vision impairment.

The instructions were in three formats: visual, tactile, and auditory. The instructions consisted of text broken into sections of explanation, graphing questions, and multiple choice answers. Examples of the instructions can be found in Appendix X¹⁰ and in Figure 48. The visual instructions were in 24¹¹-point font. Some of the explanation sections took more than one page. Every graphing question and answers for that particular question were together on a single page, and separated from all other text. The tactile instructions were in Grade 1 braille.

In the audio option, the text instructions or question and answers were heard. The speech was a prerecorded human voice. Once a section of text (a graphing problem or piece of the instructions) was read, the user had the option to repeat the question, move to the previous section, or move to the next question. The speech was embedded as part of the GNIE program. If users were not using the GNIE graph, the graph's speech and non-speech audio was disabled, and the user only heard the instructions. If users were using the GNIE graph, the graph's audio was enabled.

Each set of instructions had four parts. The first and third parts were training. They were almost identical, but one was auditory graph training and the other was tactile or visual graph training on the mathematics topics covered. At the end of each training was a short test of five questions, where participants could review the answers and ask questions. The second and fourth parts were test questions and possible answers. Each of these were

¹⁰Add to appendix and mark here

¹¹verify size

Figure 48: The instructions for the test were in three formats. From left, the images show visual graphs, tactile graphics, and auditory graphs.

Figure 49: the graphs for the tests were in three formats. From left, the images show visual graphs, tactile graphics, and auditory graphs.

in both formats, to allow for counterbalancing the graphing tests by format and order.

The graphs were in three formats: visual, tactile, and auditory. The number lines consisted of a range from -4 to 4; the graphs had the same range in both axes. The visual graphs were made with large text and high contrast lines to facilitate viewing. The tactile graphics were made a Tiger embosser, and iteratively improved, with the specialist support of the Alternative Media Access Center. The auditory graphs were generated with GNIE, but the adapted program only showed the graph panel (not the menu or text window). GNIE also showed the visual graphs, which were available to the low vision participants. Showing the visual graphs will probably only improve the scores of low vision participants, further testing the hypothesis that they will still do no better with the new tool. Examples of the graphs can be found in Appendix X¹² and in Figure 49.

All answers to the questions required selecting a choice of "A", "B", "C", or "D". Participants answered the graphing questions in three ways. Low vision participants used a pen to mark a sticky note next to the list of possible answers. The sticky note was aligned vertically, next to the list of possible answers. Participants could either write their answer (e.g. "A"), or put a mark next to their answer (e.g. a dot next to choice "A").

We were optimistic that blind participants could place a sticky dot next to their braille answer. We found, however, that the slight deviations made without visual feedback were enough to select an unintentional answer. Instead of *adding* a dot, we found that having users *remove* one of four dots led to greater success. Participants using the braille instructions had answer dots on a post-it next to the answer choices. Participants using the auditory instructions had a separate answer sheets. Each sheet had a rows of answers, one for each question on the practice and test. Examples of the answer methods can be found

¹²Add to appendix and mark here

Figure 50: The graphs for the tests were in three formats. From left, the images show visual graphs, tactile graphics, and auditory graphs.

in Appendix X¹³ and in Figure 50.

6.2.1.2 *Stimuli*

We developed sample test questions out of 11 of the 17 standards targeted. There were three reasons for this reduced count. First, three of the standards overlapped three other standards to the point where it was difficult to create a graphing problem that did not cover both. Of the 14 remaining, three had standard-unique concepts that would take more time to train the participants on. For example, one standard explores the use of open and closed points to indicate the difference between $x > 7$ and $x \geq 7$. We felt this could be done in the new format, but it was not worth the time to train participants. Finally, fewer standards opened the possibility for more time for training and other questions. We found this useful, since it turned out that the entire process took between 2-4 hours.

There were two sets of tests, each having auditory, paper, and tactile formats. Each test set had 11 questions (one from each standard), in randomized order before printing (each test was randomized in the same way, but the order between test sets was different; the instructions and answers were in the same order as the graphs). Generally speaking, the questions between tests were very similar. In some cases the questions were identical, with slight changes to the graph to make a different answer true. See Appendix X¹⁴ for the full set of graphing questions. More specifics of the stimuli used in GNIE can be found in Section ??.

6.2.1.3 *Participants*

Participants were recruited with the help of staff at the CVI and at GAB, and consisted of organization staff, clients, and students. Participation by GAB students was precluded by permission from the state, district, and parents, along with standard permission from the participants. Student participants also were required to be in a grade above sixth grade,

¹³Add to appendix and mark here

¹⁴Reference the graphs and questions appendix.

since it would then be likely that the students had already learned the graphing concepts on the test¹⁵.

6.2.1.4 Task

In general, it took participants between 1-4 hours to complete the task. Participants could stop the task at any time, and continue at a later time. This made the study more convenient for CVI staff and clients and GAB students who have several activities each day. The narrative below describes the flow from start to finish, but remember this study often occurred for a participant over several days. Participant progress and data was marked by a special index, which was the participants birth year, their initials, and their birth month. This way, participants could easily remember their number while we did not have to store participant names.

Participants were given information about the study and signed informed consent with a reader and witness. Participants were then led to a seat, next to a table with a laptop, instructions, and answers. The participant had a series of instructions and tests, based on their level of vision impairment and a counterbalance on the format of the first test and the order of the first test (test 1 or test 2 first).

Participants started with the instructions of their first format. Early on, the instructions ask for experimenter feedback (e.g. to see whether the participant knows how to find the origin), but the participant is generally independent during each part. Participants were asked to tell the instructor when they were done with each section. The instructor would then set up the participant with the next section. The four sections were: first format instructions, first format test, second format instructions, second format test. The order of the test versions and the order of the audio and non-audio (visual or tactile) formats was counterbalanced between participants. Generally, each section took between 20-40 minutes.

Each format's instructions were had about the same wording, with slightly more explanation for the auditory training. The goal of the training was to remind participants about

¹⁵While technically the students are learning from the Georgia Performance Standards (GPS), these standards are close to the CCS in terms of graphing for sixth grade, so we did not expect major gaps in previous knowledge.

the basics of graphing on number lines and graphs, so that they would be competent to answer graphing questions. As outlined in Section ??¹⁶, the major steps to learn for most of the graphing standards in CCS grade 6 are to be able to find a tick mark, find a tick mark label, find a point, find a point label, and find the origin. These were all covered, within the context of solving graphing problems.

Each test had 11 multiple choice questions. There was one graphing question for each standard used in the test. After participants completed the test, they would continue onto the next format, completing the training and subsequent test. Participants ended with a survey, and payment of \$15 per hour of their time¹⁷.

6.2.2 Pilot Study

Pilot studies were conducted in February and April 2012, as part of the classroom sessions described in Section 6.1.2. For the pilot studies, about 10 students were asked reading and writing questions, and low vision students were given visual printouts that were easier to read. The students were given 30 minutes to solve the questions. Speed and accuracy of the answers was recorded, along with student feedback. Eight standards were evaluated.

The first study established reasonable expectations for performance in the lab evaluations, identified classroom issues with using the tools, and piloted the materials for visually impaired students. In a classroom setting, about 10 visually impaired students solved validated graph questions in these formats: enlarged visual graphs and/or tactile graphics, and with GNIE. This was enough to establish that GNIE is performing sufficiently well to evaluate with more rigorous methods.

In a group setting, students were given 30 minutes to solve the questions (this was part of the earlier classroom sessions described in Section 6.1.2). Speed and accuracy of the answers was recorded, along with student feedback. This evaluation had two iterations, evaluating eight of the candidate standards (Graphing Standards 2,5,6,7,9,10,11, and 13; the specific questions for the second iteration are available in Appendix B).

¹⁶Get a labeled section for the overview of standard steps.

¹⁷Participants who took more time completing the examination were paid more money.

6.2.2.1 Pilot Results

Data analysis is continuing, but the initial results are promising. Students were able to quickly read and write points with both paper and electronic formats. Images showing reading and writing accuracy are in Figures 51 - 54.

One blind student had the same level of accuracy for four writing problems analyzed. Furthermore, the incorrect answers were incorrect for each format on the same standard, indicating that the student had difficulty understanding the standard, instead of difficulties using the tool. For the four writing problems, the student had a mean time of 2.13 minutes for audio and 2.05 minutes for tactile, with a median time of 1.55 minutes for audio and 2.00 minutes for tactile.

6.2.2.2 Expected Outcomes

The expected outcomes are driven by a single principle of fairness: blind students should be allowed to use a testing accommodation that enables them to solve the problem in a similar conceptual manner as their sighted and low vision peers. Out of this principle comes four hypotheses:

H_1 For visual graph users, graph test performance on auditory graphs is less than one point above the average of visual graphs.

H_2 For tactile graph users, graph test performance on auditory graphs is greater than two points below the average of visual graphs.

H_3 The average score for a question using the visual graph will be significantly correlated to the average score for a question using the auditory graph, with a regression line slope near 1 and intercept near 0.

H_4 The average score for a question using the tactile graph will be significantly correlated to the average score for a question using the auditory graph, with a regression line slope near 1 and intercept near 0.

Hypotheses 1 and 2 stem from the accommodations literature (discussed in more detail in Section ??). Essentially, if the accommodation is fair, the blind student should be doing no

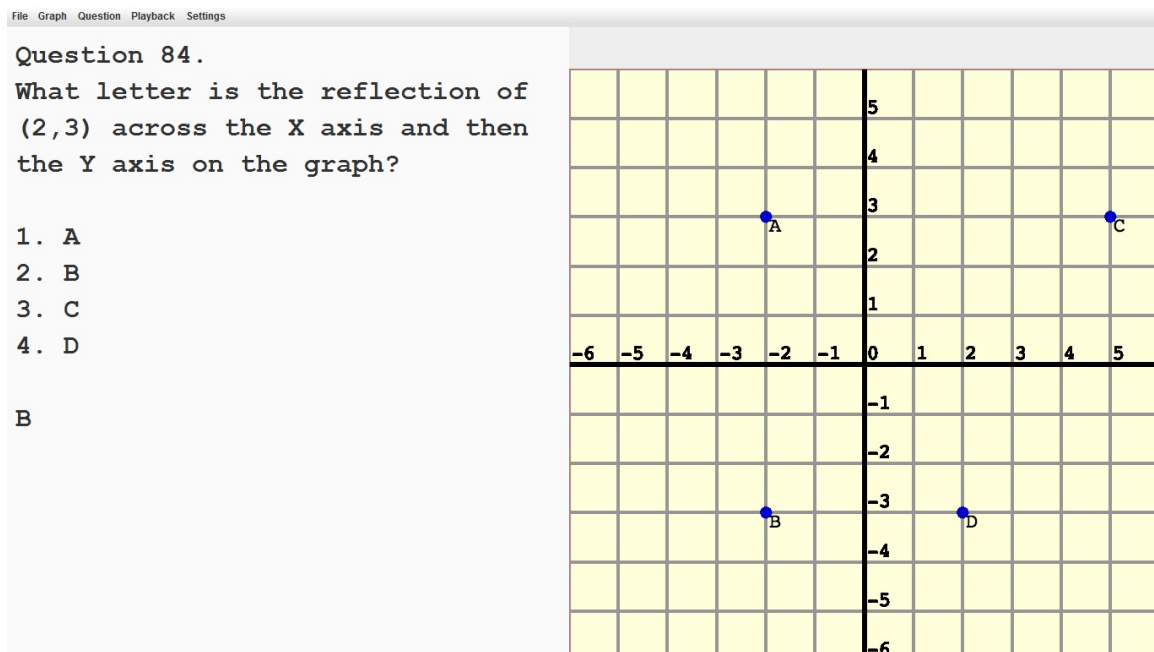


Figure 51: Identifying the point that is a reflection of $(2,3)$ across the X and Y axes. Graphing Standard 7.

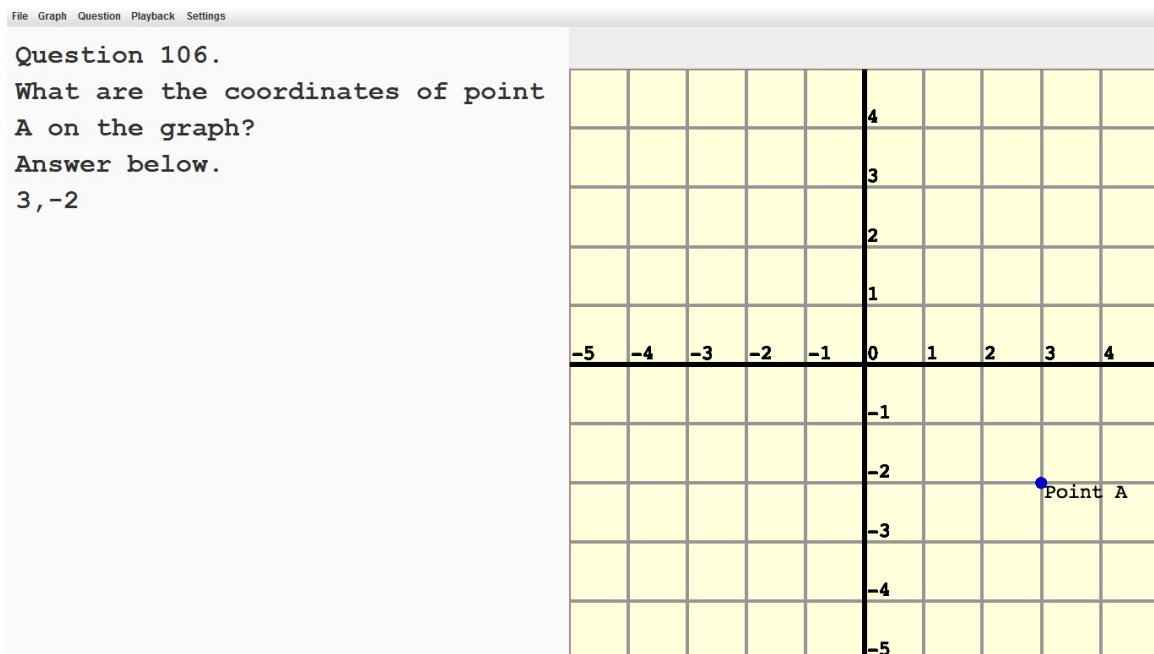


Figure 52: Finding the coordinates of point A. Graphing Standard 9.

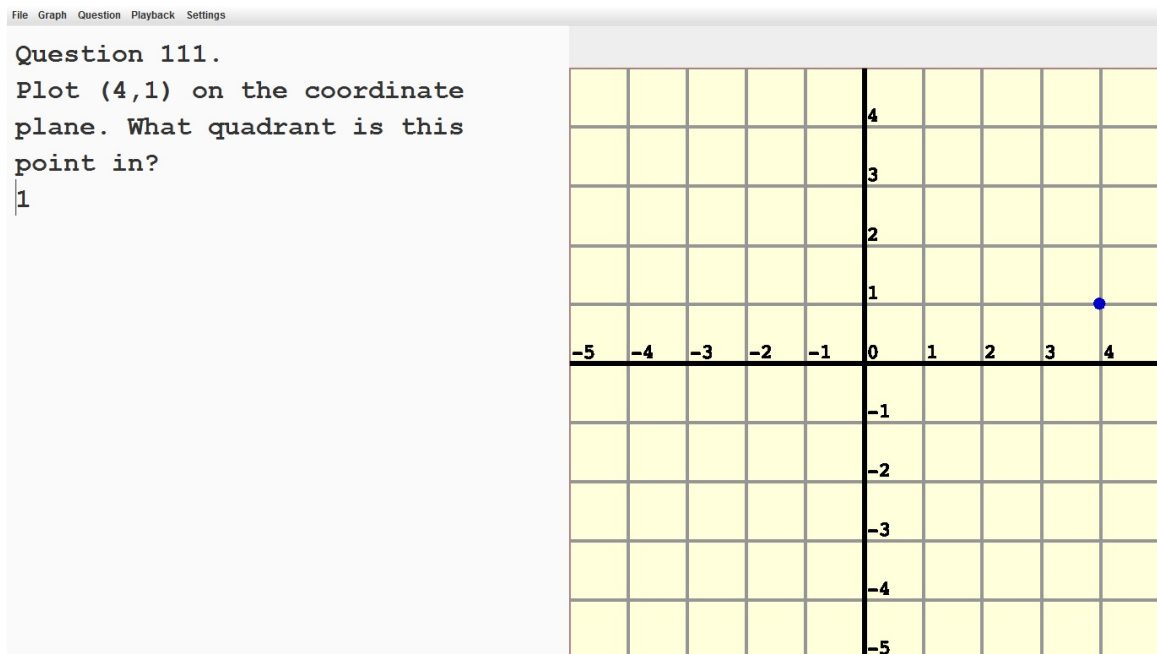


Figure 53: Plotting the point $(4,1)$ on the coordinate plane and identify the quadrant. Graphing Standard 6.

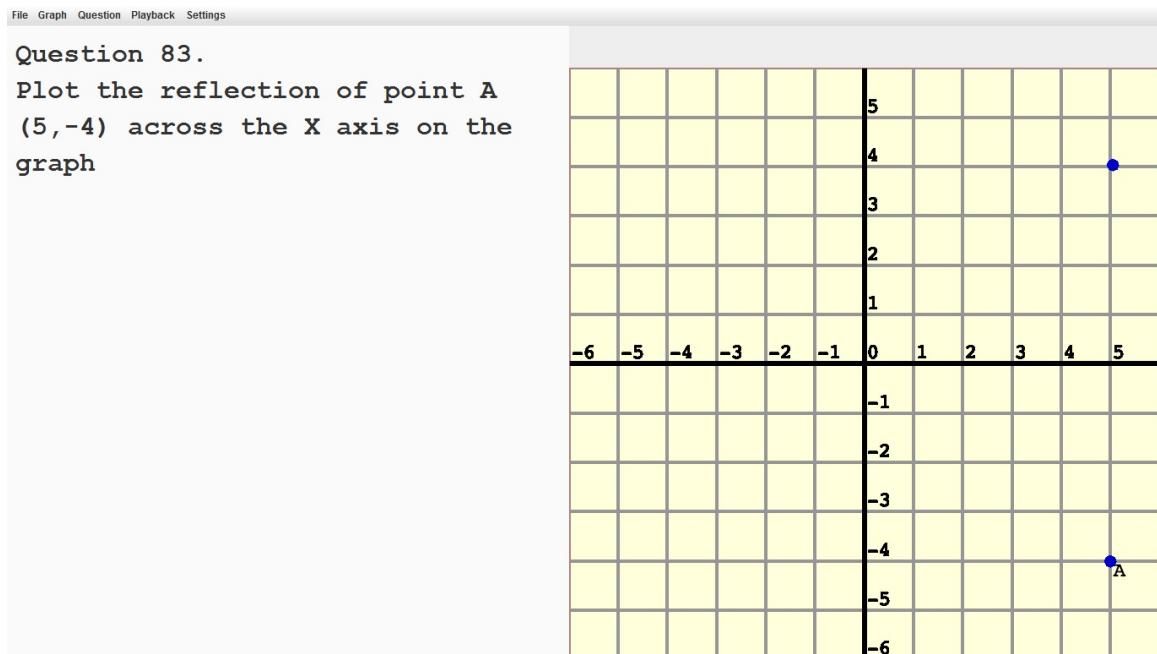


Figure 54: Plotting the reflection of $(5, -4)$ over the x axis. Graphing Standard 7.

worse, and the sighted and low vision student should not show significant performance gains. The magnitudes of the differences (1 and 2) were selected to reflect important differences in point scores. The lower bound for the tactile graphs case was selected to be 2 below as a conservative check; auditory graph scores may be slightly lower due to inexperience, so an extra point was added to the bounds. Since there were 11 graphing questions, the percent bounds were 9.09% over and 18.18% under the mean.

Hypotheses 3 and 4 are novel to this experiment, but, I believe, are often overlooked. It is likely that some kinds of graphing questions are more difficult than others on the test. If they are difficult due to the concept, then they should be difficult regardless of the perceptual format. A keystone assumption of this thesis (and accessible graphs) is that graphs can maintain their conceptual integrity through transformations into auditory and tactile formats¹⁸. Therefore, any inherent difficulty in a graphing problem should maintain itself in a transformation. The correlation of 1 means that when a more difficult problem occurs, it is equally more difficult in both formats. While H_1 and H_2 are more behavioral, H_3 and H_4 suggest a more cognitive connection between the two formats being tested.

With wording like "near 1", H_3 and H_4 risk becoming difficult-to-test hypotheses. To clarify the test, the null hypothesis would be: The average score for a question using the (visual or tactile) graph will not be significantly correlated to the average score for a question using the auditory graph, and/or the regression line of the correlation will not have a slope with a confidence interval containing 1, nor an intercept with a confidence interval containing 0.

Since there were two tests (one in each format) for each participant, their scores could be compared by averages between the formats. In addition, the standards-questions could be compared by whether the tactile version average was correlated with the auditory version average, as well as the visual version average and the auditory version average.

One final note. There could be many more tests conducted to narrow down the nature of the relationship between the scores. However, since this research is in accessibility, and

¹⁸Furthermore, the concept of the graph can be un-grounded from the visual graph. For example, the acAGM uses a perception-independent model. Visual, tactile, and auditory graphs can be first-class displays, not depending on the other.

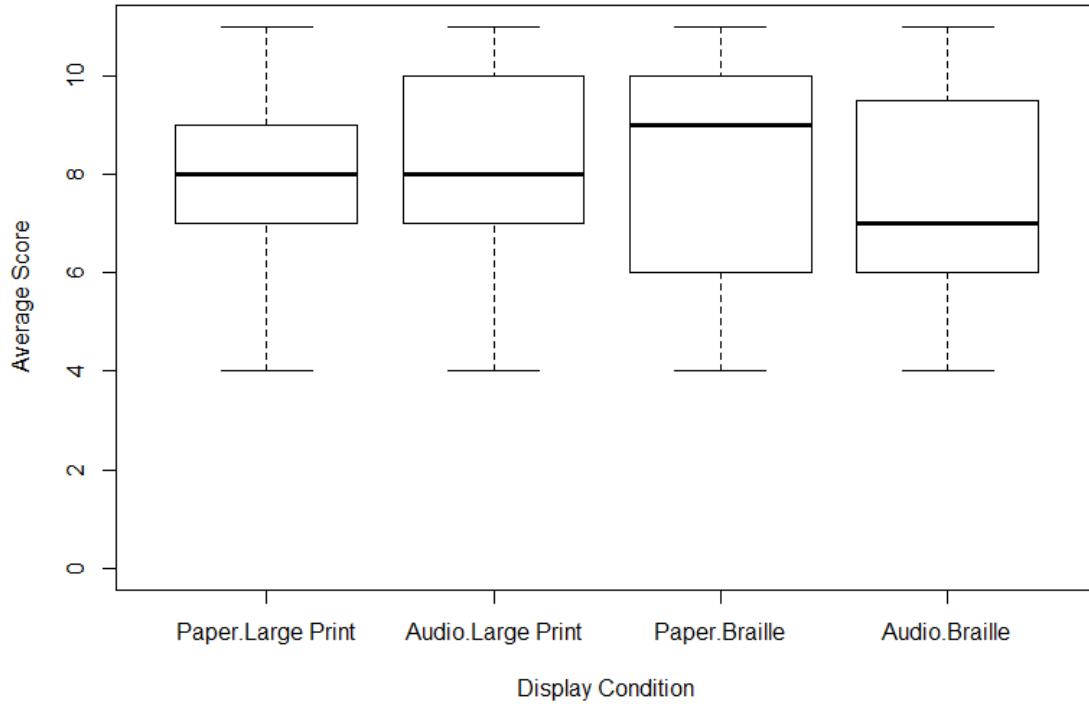


Figure 55: Test scores by display condition.

it is difficult to find many participants, every effort has been made to minimize the number of statistical tests.

6.2.3 Results

Tests were evaluated with an familywise alpha of $\alpha = 0.050/2 = .025$, using the conservative Bonferroni correction.

6.2.3.1 Low vision users perform no better with auditory graphs

The scores for low vision participants was 8.19 (SD 1.69) for visual graphs, and 8.00 (SD 2.04) for auditory graphs. A t-test confirmed that the auditory graphs are significantly lower than 1 above the average, $m_{audio} < 9.19$, $t(35) = -3.51$, $p < 0.025$. Figure 55 shows a box plot of large print and braille cases, compared with audio. When using GNIE for the test, low vision participants have, on average, a score lower than one point above their large print score.

6.2.3.2 *Blind users perform no worse with auditory graphs*

The scores for blind participants was 8.20 (SD 2.40) for tactile graphics, and 7.60 (SD 1.98) for auditory graphs. A t-test confirmed that the auditory graphs are significantly higher than 2 points below the average, $m_{audio} > 6.20$, $t(19) = 3.16$, $p < 0.025$. When using GNIE for the test, blind participants have, on average, a score higher than two points below their tactile graphics score.

6.2.3.3 *Problem difficulty is similar between visual and auditory graphs*

There was a significant correlation between the average score for each standard in the large print test to the average score for each standard in the auditory graph test for low vision participants, $F(1, 9) = 41.42$, $p < 0.025$, $R^2 = 0.8215$. The line of best fit was $y = 0.827x + 0.112$, which had a confidence interval (for the formula $y = mx + b$) of $0.536 < x < 1.12$ and $-0.114 < b < 0.337$. Figure 56 shows a plot of the relationship. The average score for each standard is correlated between formats for low vision participants.

6.2.3.4 *Problem difficulty is similar between tactile and auditory graphs*

There was a significant correlation between the average score for each standard in the braille test to the average score for each standard in the auditory graphs test for blind participants, $F(1, 9) = 7.493$, $p < 0.025$, $R^2 = 0.454$. The line of best fit was $y = 0.944x - 0.126$, which had a confidence interval of $0.164 < x < 1.172$ and $-0.600 < b < 0.575$. Figure 57 shows a plot of the relationship. The average score for each standard is correlated between formats for blind participants.

While there was a correlation, the R^2 was much lower than in the low vision test, describing less than half of the observed variance. Standards 5 and 7 had questions that asked participants to flip a point over the origin or one of the axes. The results showed that these problems were particularly challenging for blind users using GNIE. This may be a specific limitation of the tool to these types of questions. A post-hoc analysis showed that without these points, a much stronger correlation occurs, $F(1, 7) = 22.17$, $p < 0.01$, $R^2 = 0.760$, $y = 0.881x + 0.0837$. Figure 58 shows a plot of the relationship.

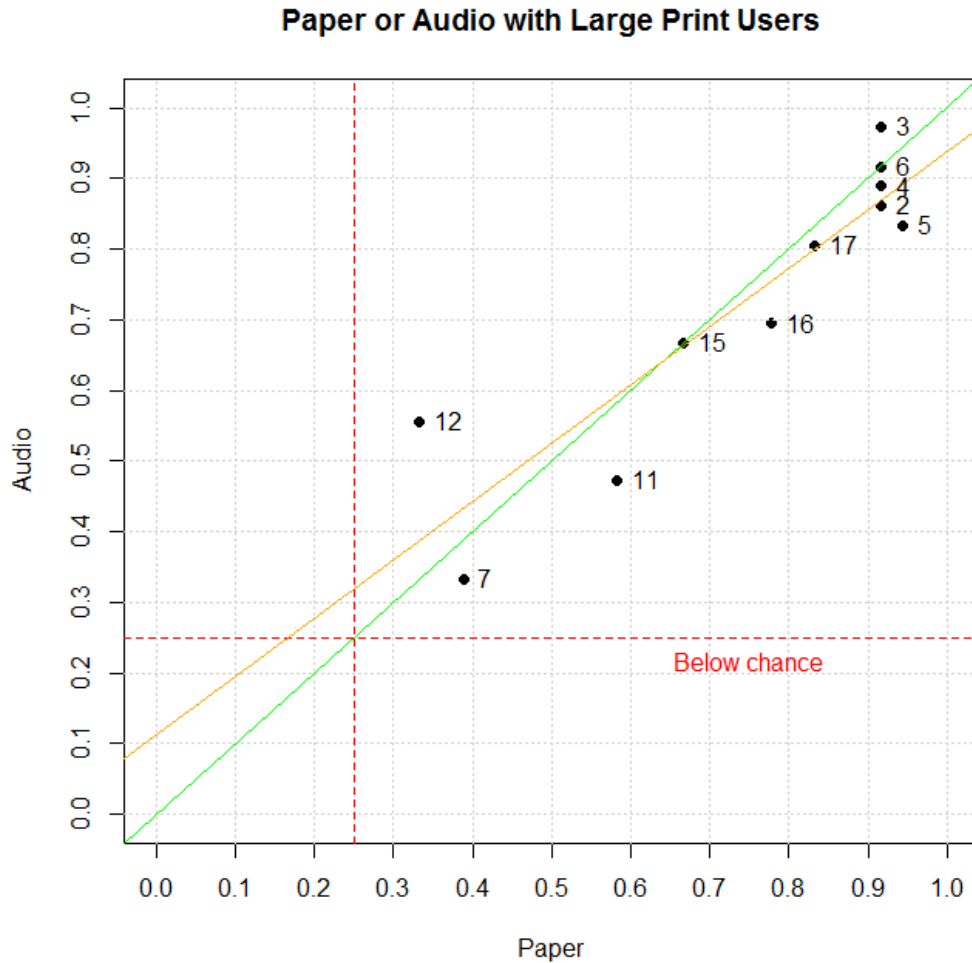


Figure 56: Test score averages for each standard, comparing large print ("paper") to the GNIE ("audio") scores. The center green line is $y=x$, the theoretical fit line for the scores.

6.2.3.5 Other findings

We found a two major problems in our initial presentation for the blind users. First, for most of them, the grade 1 braille¹⁹ presented was not ideal. Many adults had not learned braille, or knew the very basics. Others knew braille, but used grade 2 braille much more often²⁰, so they found reading Grade 1 to be more time consuming. For both groups,

¹⁹Grade 1 braille is a direct mapping of letters and numbers to braille characters. Braille has no differences in grammar to American English. The result is Grade 1 braille is like Morse Code: a simple mapping of a visible character to a tactile character (and a visible character, since braille can be read by sight).

²⁰Grade 2 braille uses contractions, so it is more efficient for reading.

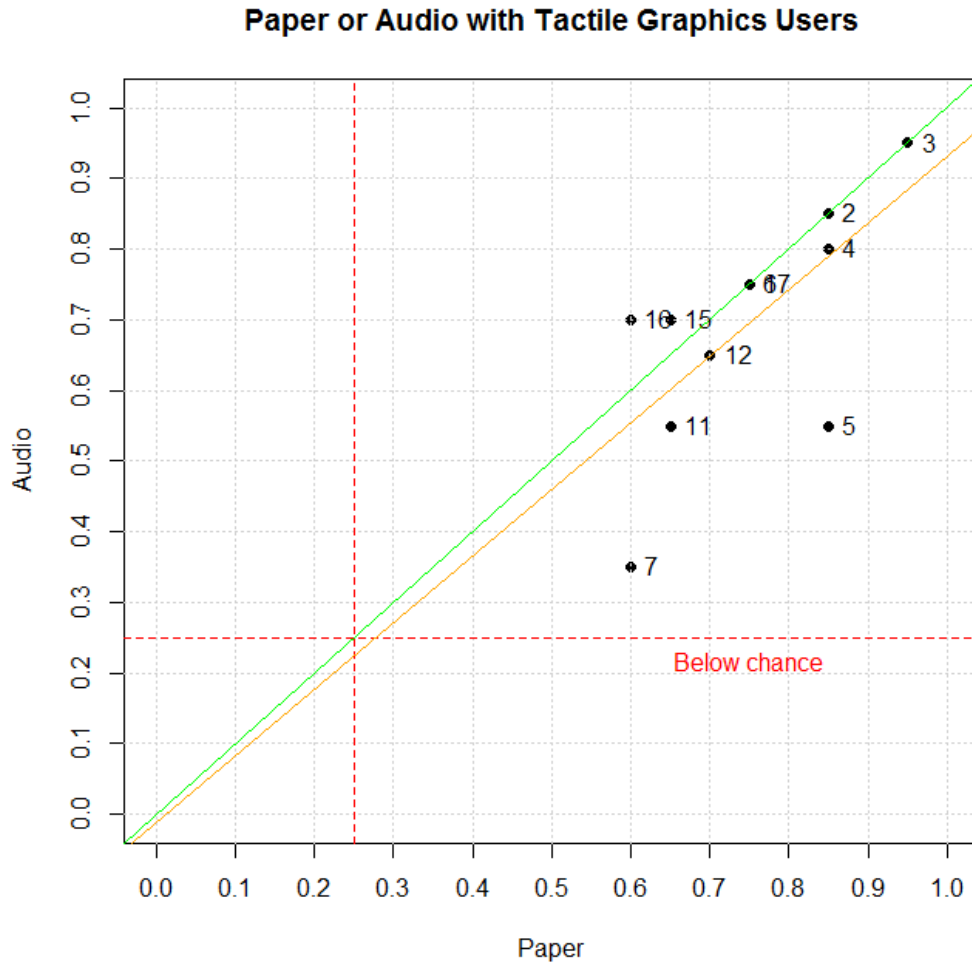


Figure 57: Test score averages for each standard, comparing braille ("paper") to the GNIE ("audio") scores. The center green line is $y=x$, the theoretical fit line for the scores.

an audio option for the instructions was often more helpful than the braille instructions provided.

6.2.4 Discussion

Considering the limited interaction participants have had with auditory graphs, they performed remarkably well with GNIE. A key consideration is that the tests were designed to ask participants about their knowledge of graphing. If they knew how to graph, they could use the tool (in whatever format) to answer their question.

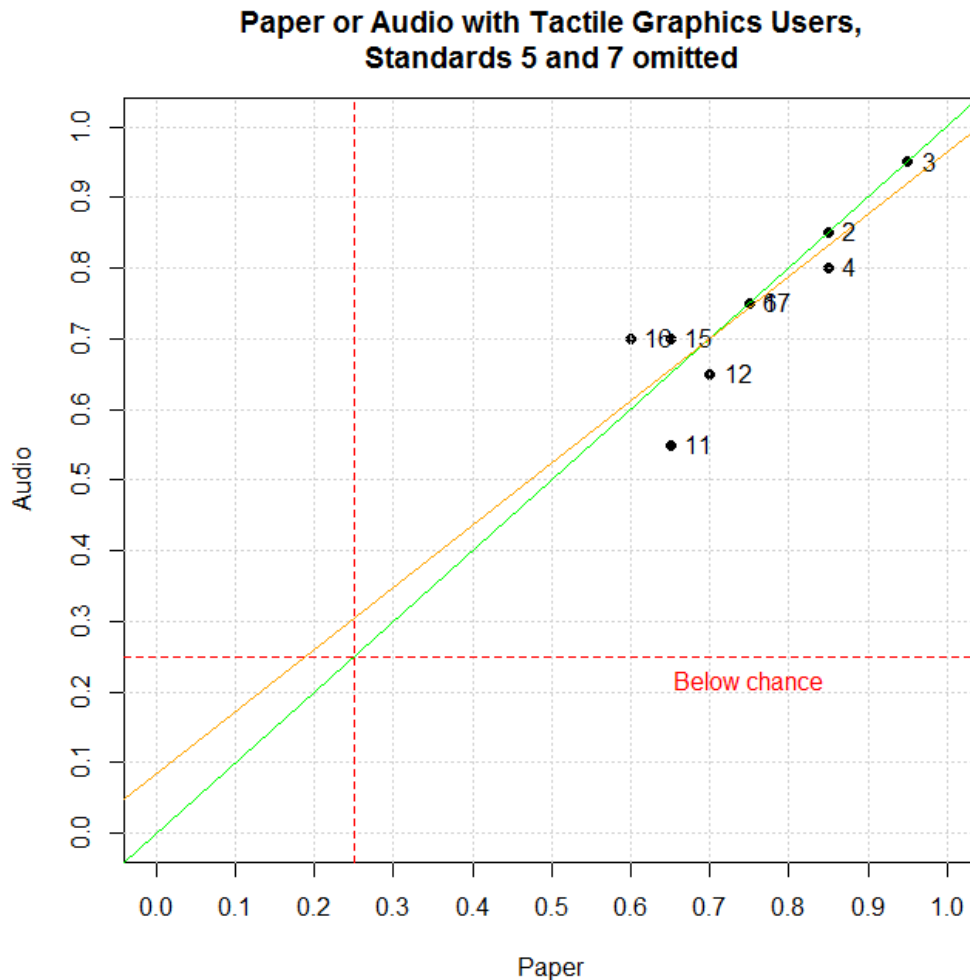


Figure 58: Test score averages for each standard, comparing braille ("paper") to the GNIE ("audio") scores, and omitting questions for standards 5 and 7. The center green line is $y=x$, the theoretical fit line for the scores.

It appears that GNIE scores could be further improved. Figures 59 and ?? show histograms of the low vision and blind user's scores. The distribution of scores for low vision users who use paper appears to follow a normal curve, peaking at a score of 8. Somewhat similarly, the distribution of scores for blind users who use tactile graphics also follows a steep curve, peaking at 10, if the six low scores are considered anomalies²¹. The GNIE scores in both cases, however, are relatively flat. We expect that there is a learning effect occurring, where users who have caught on to GNIE are performing near their potential,

²¹It is likely that these low scores are due to people not familiar with tactile graphics, similar to the familiarity problem with GNIE.

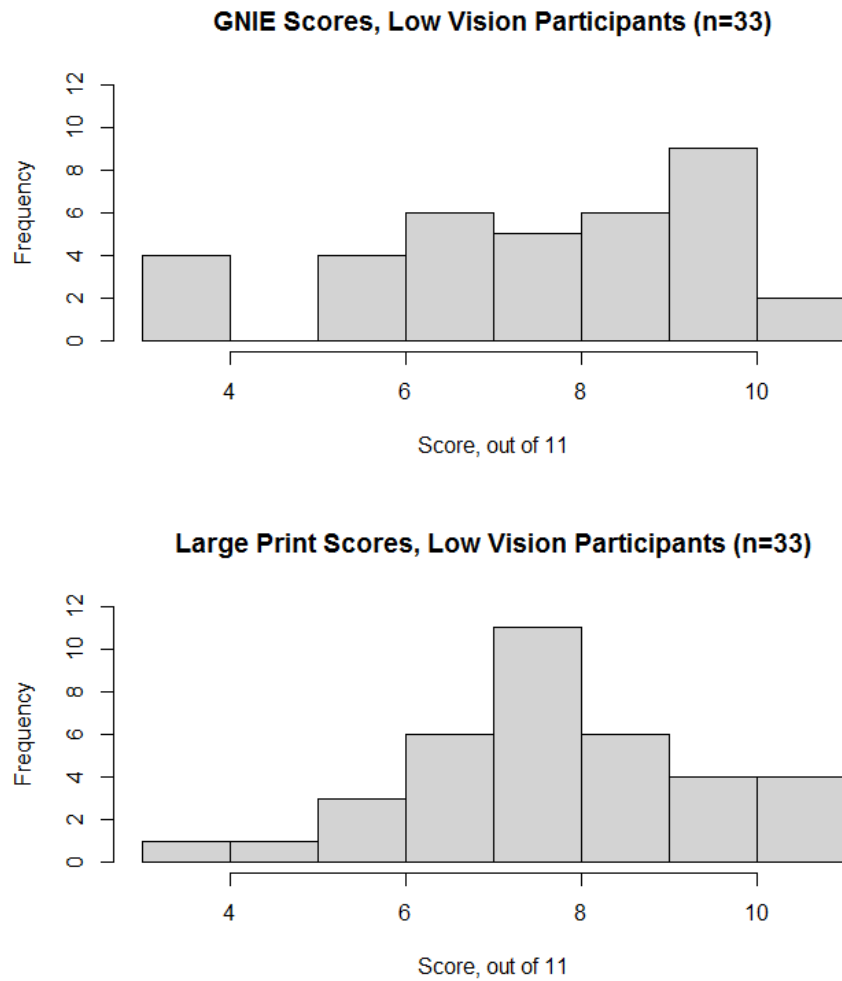


Figure 59: Histograms for the scores from low vision users, showing large print (top) and GNIE (bottom) conditions.

while others may still be learning how to use the system. In addition, in some cases users may be *more* comfortable with the new system (independent of any real, broad benefit to their user population). These participants may score better because GNIE provides benefits beyond their current tools, because the participants prefer computers, because the test they took on the computer was slightly easier, or a number of other reasons.

This examination study went further than most evaluations of assistive technologies for education. It met all four challenging goals, and arguably is a viable testing accommodation. There are three major limitations to the results, however. First, other aspects of the test were not evaluated. While not central to this thesis, issues such as text-to-speech engine

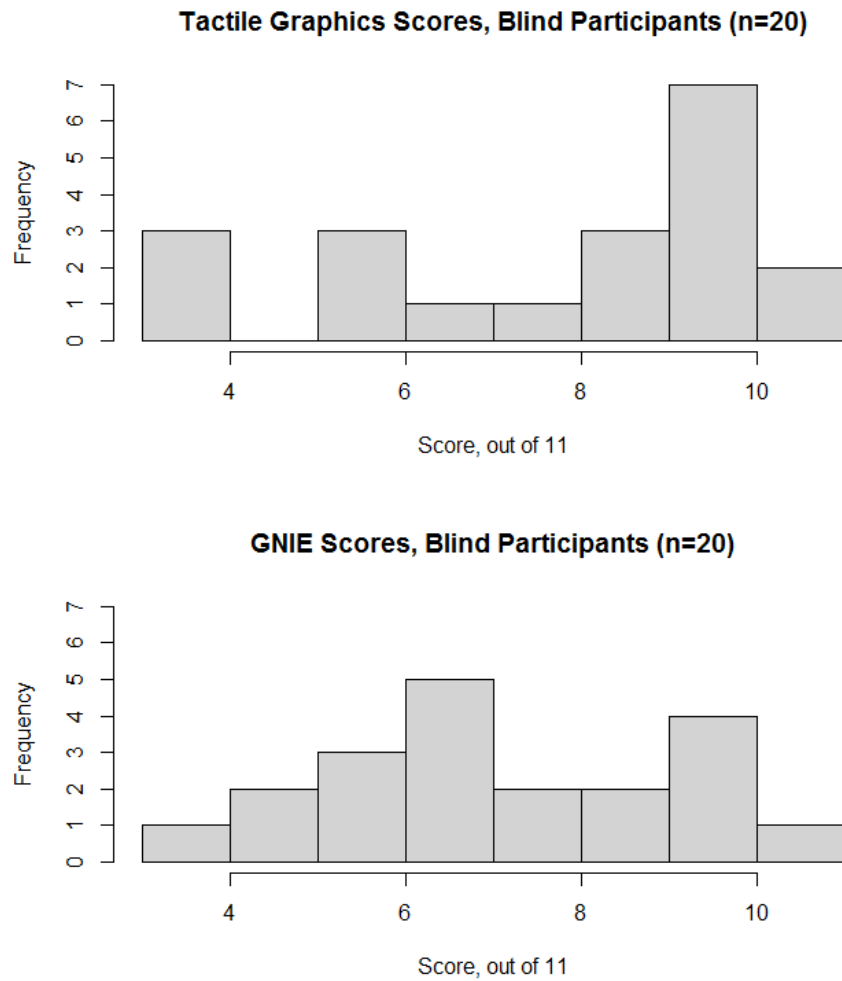


Figure 60: Histograms for the scores from tactile graphics users, showing tactile (top) and GNIE (bottom) conditions.

used and even the type of keyboard provided are key components to blind student successes in testing. Second, this test only evaluated reading questions. This is primarily a limitation of the method, which performs better when there are highly limited options for results (e.g. multiple choice questions). Evaluating the boundaries of when a student does or does not put a point exactly on a line are more ambiguous. Nevertheless, an expanded study on evaluating performance with written questions would compliment this study, and could be modeled after the same method. Finally, the number of graphing questions was limited to two, for only 11 standards. A more in-depth look at fewer standards, and with more students, would provide more statistical power for exploring the relationship between

performance on visual, tactile, and auditory graphs.

6.3 General Discussion

Auditory graphs can be used in classroom situations and on tests. The first section of this chapter showed how graphs work in the classroom, and that students can solve graphing questions with visual graphs, auditory graphs, or tactile graphics, within the scope of a complex classroom situation. Students learned the basics of how to use graphs through the Navy program, and were often competent in the basics within 10-15 minutes of training and games. Students reported approximately equal opinions of auditory graphs compared to their counterparts, but a few had troubles making the new tool work well for their situation. Teachers reported that they could make graphs for GNIE in about the same time as it took them to make visual and/or tactile graphics.

In the test experiment, low vision and blind participants showed that auditory graphs are a suitable testing accommodation. Specifically, the graphs in GNIE allow blind participants to perform near their potential with tactile graphics, yet do not give an advantage, since low vision participants did not outperform their visual graph scores.

The process outlined here was applied to GNIE, but can be used for many different assistive technologies for education. Often, the questions of functionality, appropriateness, and fairness are raised with testing accommodations; the data-driven approach laid out in this chapter provides a way to more fully understand the potential of any educational technology, particularly for accessibility.

Bibliography

APPENDIX A

COMMON CORE STANDARDS

This chapter includes the Common Core Standards for Mathematics, Grade 6. This is taken directly from [?], available at <http://corestandards.org>.

Mathematics | Grade 6

In Grade 6, instructional time should focus on four critical areas: (1) connecting ratio and rate to whole number multiplication and division and using concepts of ratio and rate to solve problems; (2) completing understanding of division of fractions and extending the notion of number to the system of rational numbers, which includes negative numbers; (3) writing, interpreting, and using expressions and equations; and (4) developing understanding of statistical thinking.

(1) Students use reasoning about multiplication and division to solve ratio and rate problems about quantities. By viewing equivalent ratios and rates as deriving from, and extending, pairs of rows (or columns) in the multiplication table, and by analyzing simple drawings that indicate the relative size of quantities, students connect their understanding of multiplication and division with ratios and rates. Thus students expand the scope of problems for which they can use multiplication and division to solve problems, and they connect ratios and fractions. Students solve a wide variety of problems involving ratios and rates.

(2) Students use the meaning of fractions, the meanings of multiplication and division, and the relationship between multiplication and division to understand and explain why the procedures for dividing fractions make sense. Students use these operations to solve problems. Students extend their previous understandings of number and the ordering of numbers to the full system of rational numbers, which includes negative rational numbers, and in particular negative integers. They reason about the order and absolute value of rational numbers and about the location of points in all four quadrants of the coordinate plane.

(3) Students understand the use of variables in mathematical expressions. They write expressions and equations that correspond to given situations, evaluate expressions, and use expressions and formulas to solve problems. Students understand that expressions in different forms can be equivalent, and they use the properties of operations to rewrite expressions in equivalent forms. Students know that the solutions of an equation are the values of the variables that make the equation true. Students use properties of operations and the idea of maintaining the equality of both sides of an equation to solve simple one-step equations. Students construct and analyze tables, such as tables of quantities that are in equivalent ratios, and they use equations (such as $3x = y$) to describe relationships between quantities.

(4) Building on and reinforcing their understanding of number, students begin to develop their ability to think statistically. Students recognize that a data distribution may not have a definite center and that different ways to measure center yield different values. The median measures center in the sense that it is roughly the middle value. The mean measures center in the sense that it is the value that each data point would take on if the total of the data values were redistributed equally, and also in the sense that it is a balance point. Students recognize that a measure of variability (interquartile range or mean absolute deviation) can also be useful for summarizing data because two very different sets of data can have the same mean and

median yet be distinguished by their variability. Students learn to describe and summarize numerical data sets, identifying clusters, peaks, gaps, and symmetry, considering the context in which the data were collected.

Students in Grade 6 also build on their work with area in elementary school by reasoning about relationships among shapes to determine area, surface area, and volume. They find areas of right triangles, other triangles, and special quadrilaterals by decomposing these shapes, rearranging or removing pieces, and relating the shapes to rectangles. Using these methods, students discuss, develop, and justify formulas for areas of triangles and parallelograms. Students find areas of polygons and surface areas of prisms and pyramids by decomposing them into pieces whose area they can determine. They reason about right rectangular prisms with fractional side lengths to extend formulas for the volume of a right rectangular prism to fractional side lengths. They prepare for work on scale drawings and constructions in Grade 7 by drawing polygons in the coordinate plane.

Grade 6 Overview

Ratios and Proportional Relationships

- Understand ratio concepts and use ratio reasoning to solve problems.

The Number System

- Apply and extend previous understandings of multiplication and division to divide fractions by fractions.
- Compute fluently with multi-digit numbers and find common factors and multiples.
- Apply and extend previous understandings of numbers to the system of rational numbers.

Expressions and Equations

- Apply and extend previous understandings of arithmetic to algebraic expressions.
- Reason about and solve one-variable equations and inequalities.
- Represent and analyze quantitative relationships between dependent and independent variables.

Geometry

- Solve real-world and mathematical problems involving area, surface area, and volume.

Statistics and Probability

- Develop understanding of statistical variability.
- Summarize and describe distributions.

Mathematical Practices

1. Make sense of problems and persevere in solving them.
2. Reason abstractly and quantitatively.
3. Construct viable arguments and critique the reasoning of others.
4. Model with mathematics.
5. Use appropriate tools strategically.
6. Attend to precision.
7. Look for and make use of structure.
8. Look for and express regularity in repeated reasoning.

Ratios and Proportional Relationships**6.RP****Understand ratio concepts and use ratio reasoning to solve problems.**

1. Understand the concept of a ratio and use ratio language to describe a ratio relationship between two quantities. *For example, "The ratio of wings to beaks in the bird house at the zoo was 2:1, because for every 2 wings there was 1 beak." "For every vote candidate A received, candidate C received nearly three votes."*
2. Understand the concept of a unit rate a/b associated with a ratio $a:b$ with $b \neq 0$, and use rate language in the context of a ratio relationship. *For example, "This recipe has a ratio of 3 cups of flour to 4 cups of sugar, so there is $3/4$ cup of flour for each cup of sugar." "We paid \$75 for 15 hamburgers, which is a rate of \$5 per hamburger."*¹
3. Use ratio and rate reasoning to solve real-world and mathematical problems, e.g., by reasoning about tables of equivalent ratios, tape diagrams, double number line diagrams, or equations.
 - a. Make tables of equivalent ratios relating quantities with whole-number measurements, find missing values in the tables, and plot the pairs of values on the coordinate plane. Use tables to compare ratios.
 - b. Solve unit rate problems including those involving unit pricing and constant speed. *For example, if it took 7 hours to mow 4 lawns, then at that rate, how many lawns could be mowed in 35 hours? At what rate were lawns being mowed?*
 - c. Find a percent of a quantity as a rate per 100 (e.g., 30% of a quantity means $30/100$ times the quantity); solve problems involving finding the whole, given a part and the percent.
 - d. Use ratio reasoning to convert measurement units; manipulate and transform units appropriately when multiplying or dividing quantities.

The Number System**6.NS****Apply and extend previous understandings of multiplication and division to divide fractions by fractions.**

1. Interpret and compute quotients of fractions, and solve word problems involving division of fractions by fractions, e.g., by using visual fraction models and equations to represent the problem. *For example, create a story context for $(2/3) \div (3/4)$ and use a visual fraction model to show the quotient; use the relationship between multiplication and division to explain that $(2/3) \div (3/4) = 8/9$ because $3/4$ of $8/9$ is $2/3$. (In general, $(a/b) \div (c/d) = ad/bc$.) How much chocolate will each person get if 3 people share $1/2$ lb of chocolate equally? How many $3/4$ -cup servings are in $2/3$ of a cup of yogurt? How wide is a rectangular strip of land with length $3/4$ mi and area $1/2$ square mi?*

Compute fluently with multi-digit numbers and find common factors and multiples.

2. Fluently divide multi-digit numbers using the standard algorithm.
3. Fluently add, subtract, multiply, and divide multi-digit decimals using the standard algorithm for each operation.
4. Find the greatest common factor of two whole numbers less than or equal to 100 and the least common multiple of two whole numbers less than or equal to 12. Use the distributive property to express a sum of two whole numbers 1–100 with a common factor as a multiple of a sum of two whole numbers with no common factor. *For example, express $36 + 8$ as $4(9 + 2)$.*

¹Expectations for unit rates in this grade are limited to non-complex fractions.

Apply and extend previous understandings of numbers to the system of rational numbers.

5. Understand that positive and negative numbers are used together to describe quantities having opposite directions or values (e.g., temperature above/below zero, elevation above/below sea level, credits/debits, positive/negative electric charge); use positive and negative numbers to represent quantities in real-world contexts, explaining the meaning of 0 in each situation.
6. Understand a rational number as a point on the number line. Extend number line diagrams and coordinate axes familiar from previous grades to represent points on the line and in the plane with negative number coordinates.
 - a. Recognize opposite signs of numbers as indicating locations on opposite sides of 0 on the number line; recognize that the opposite of the opposite of a number is the number itself, e.g., $-(-3) = 3$, and that 0 is its own opposite.
 - b. Understand signs of numbers in ordered pairs as indicating locations in quadrants of the coordinate plane; recognize that when two ordered pairs differ only by signs, the locations of the points are related by reflections across one or both axes.
 - c. Find and position integers and other rational numbers on a horizontal or vertical number line diagram; find and position pairs of integers and other rational numbers on a coordinate plane.
7. Understand ordering and absolute value of rational numbers.
 - a. Interpret statements of inequality as statements about the relative position of two numbers on a number line diagram. *For example, interpret $-3 > -7$ as a statement that -3 is located to the right of -7 on a number line oriented from left to right.*
 - b. Write, interpret, and explain statements of order for rational numbers in real-world contexts. *For example, write $-3^{\circ}\text{C} > -7^{\circ}\text{C}$ to express the fact that -3°C is warmer than -7°C .*
 - c. Understand the absolute value of a rational number as its distance from 0 on the number line; interpret absolute value as magnitude for a positive or negative quantity in a real-world situation. *For example, for an account balance of -30 dollars, write $|-30| = 30$ to describe the size of the debt in dollars.*
 - d. Distinguish comparisons of absolute value from statements about order. *For example, recognize that an account balance less than -30 dollars represents a debt greater than 30 dollars.*
8. Solve real-world and mathematical problems by graphing points in all four quadrants of the coordinate plane. Include use of coordinates and absolute value to find distances between points with the same first coordinate or the same second coordinate.

Expressions and Equations

6.EE

Apply and extend previous understandings of arithmetic to algebraic expressions.

1. Write and evaluate numerical expressions involving whole-number exponents.
2. Write, read, and evaluate expressions in which letters stand for numbers.
 - a. Write expressions that record operations with numbers and with letters standing for numbers. *For example, express the calculation "Subtract y from 5" as $5 - y$.*

- b. Identify parts of an expression using mathematical terms (sum, term, product, factor, quotient, coefficient); view one or more parts of an expression as a single entity. *For example, describe the expression $2(8 + 7)$ as a product of two factors; view $(8 + 7)$ as both a single entity and a sum of two terms.*
 - c. Evaluate expressions at specific values of their variables. Include expressions that arise from formulas used in real-world problems. Perform arithmetic operations, including those involving whole-number exponents, in the conventional order when there are no parentheses to specify a particular order (Order of Operations). *For example, use the formulas $V = s^3$ and $A = 6s^2$ to find the volume and surface area of a cube with sides of length $s = \frac{1}{2}$.*
3. Apply the properties of operations to generate equivalent expressions. *For example, apply the distributive property to the expression $3(2 + x)$ to produce the equivalent expression $6 + 3x$; apply the distributive property to the expression $24x + 18y$ to produce the equivalent expression $6(4x + 3y)$; apply properties of operations to $y + y + y$ to produce the equivalent expression $3y$.*
 4. Identify when two expressions are equivalent (i.e., when the two expressions name the same number regardless of which value is substituted into them). *For example, the expressions $y + y + y$ and $3y$ are equivalent because they name the same number regardless of which number y stands for.*

Reason about and solve one-variable equations and inequalities.

5. Understand solving an equation or inequality as a process of answering a question: which values from a specified set, if any, make the equation or inequality true? Use substitution to determine whether a given number in a specified set makes an equation or inequality true.
6. Use variables to represent numbers and write expressions when solving a real-world or mathematical problem; understand that a variable can represent an unknown number, or, depending on the purpose at hand, any number in a specified set.
7. Solve real-world and mathematical problems by writing and solving equations of the form $x + p = q$ and $px = q$ for cases in which p , q and x are all nonnegative rational numbers.
8. Write an inequality of the form $x > c$ or $x < c$ to represent a constraint or condition in a real-world or mathematical problem. Recognize that inequalities of the form $x > c$ or $x < c$ have infinitely many solutions; represent solutions of such inequalities on number line diagrams.

Represent and analyze quantitative relationships between dependent and independent variables.

9. Use variables to represent two quantities in a real-world problem that change in relationship to one another; write an equation to express one quantity, thought of as the dependent variable, in terms of the other quantity, thought of as the independent variable. Analyze the relationship between the dependent and independent variables using graphs and tables, and relate these to the equation. *For example, in a problem involving motion at constant speed, list and graph ordered pairs of distances and times, and write the equation $d = 65t$ to represent the relationship between distance and time.*

Geometry

6.G

Solve real-world and mathematical problems involving area, surface area, and volume.

1. Find the area of right triangles, other triangles, special quadrilaterals, and polygons by composing into rectangles or decomposing into triangles and other shapes; apply these techniques in the context of solving real-world and mathematical problems.

2. Find the volume of a right rectangular prism with fractional edge lengths by packing it with unit cubes of the appropriate unit fraction edge lengths, and show that the volume is the same as would be found by multiplying the edge lengths of the prism. Apply the formulas $V = l w h$ and $V = b h$ to find volumes of right rectangular prisms with fractional edge lengths in the context of solving real-world and mathematical problems.
3. Draw polygons in the coordinate plane given coordinates for the vertices; use coordinates to find the length of a side joining points with the same first coordinate or the same second coordinate. Apply these techniques in the context of solving real-world and mathematical problems.
4. Represent three-dimensional figures using nets made up of rectangles and triangles, and use the nets to find the surface area of these figures. Apply these techniques in the context of solving real-world and mathematical problems.

Statistics and Probability

6.SP

Develop understanding of statistical variability.

1. Recognize a statistical question as one that anticipates variability in the data related to the question and accounts for it in the answers. *For example, "How old am I?" is not a statistical question, but "How old are the students in my school?" is a statistical question because one anticipates variability in students' ages.*
2. Understand that a set of data collected to answer a statistical question has a distribution which can be described by its center, spread, and overall shape.
3. Recognize that a measure of center for a numerical data set summarizes all of its values with a single number, while a measure of variation describes how its values vary with a single number.

Summarize and describe distributions.

4. Display numerical data in plots on a number line, including dot plots, histograms, and box plots.
5. Summarize numerical data sets in relation to their context, such as by:
 - a. Reporting the number of observations.
 - b. Describing the nature of the attribute under investigation, including how it was measured and its units of measurement.
 - c. Giving quantitative measures of center (median and/or mean) and variability (interquartile range and/or mean absolute deviation), as well as describing any overall pattern and any striking deviations from the overall pattern with reference to the context in which the data were gathered.
 - d. Relating the choice of measures of center and variability to the shape of the data distribution and the context in which the data were gathered.

APPENDIX B

GRAPHING QUESTIONS

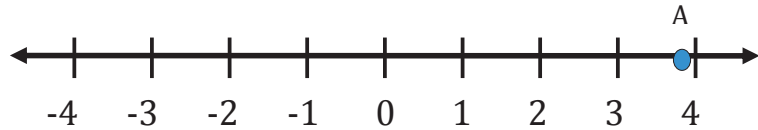
This chapter shows examples of the graphing questions evaluated with teachers. The graphing standard numbering was different during the production of these slides; For example, “CCS.6.NS.6.i” indicates the first graphing standard based on 6.NS.6. “RQ” indicates a reading question, and “WQ” indicates a writing question.

The last page shows an earlier version of a few graphing questions, in braille.

Note that these questions originally appeared as one per page. They are displayed as four per page to save space in this document.

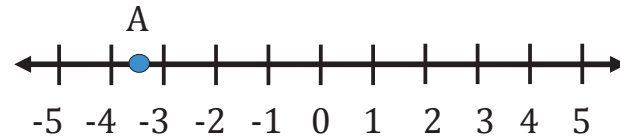
91. CCS-6.NS.6.i – Standard 1, RQ 1

Estimate the value of point A on the number line.



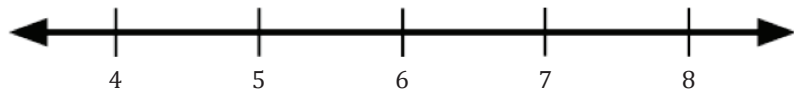
113. CCS-6.NS.6.i – Standard 1, RQ 2

Estimate the value of point A on the number line.



5. CCS-6.NS.6.i – Standard 1, WQ 1

Plot 6.3 on the number line.



66. CCS-6.NS.6.i – Standard 1, WQ 2

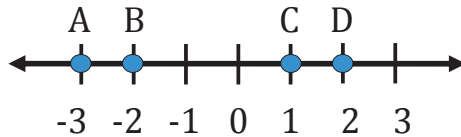
Plot 4.2 on the number line.



41. CCS-6.NS.6.a.i – Standard 2, RQ 1

What letter is the opposite of B?

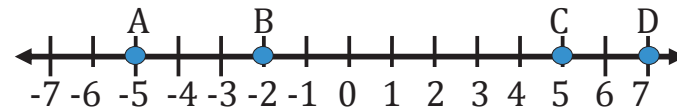
1. A
2. B
3. C
4. D



44. CCS-6.NS.6.a.i – Standard 2, RQ 2

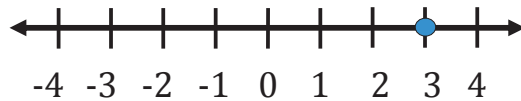
What letter is the opposite of A on the number line?

1. A
2. B
3. C
4. D



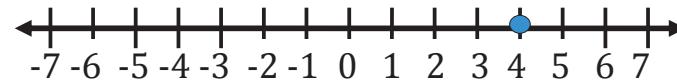
32. CCS-6.NS.6.a.i – Standard 2, WQ 1

Plot the opposite of 3 on the number line



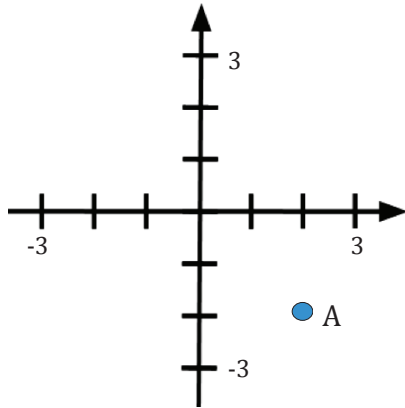
37. CCS-6.NS.6.a.i – Standard 2, WQ 2

Plot the opposite of 4 on the number line



103. CCS-6.NS.6.b.i – Standard 3, RQ 1

What quadrant is point A in on the coordinate plane?

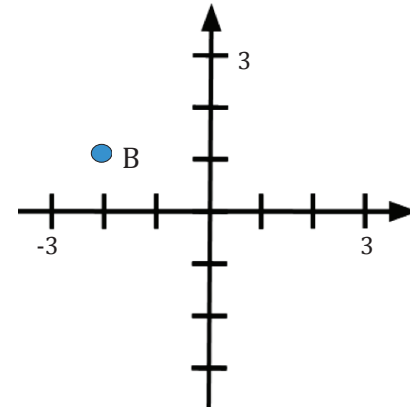


Evaluation Problems

10

104. CCS-6.NS.6.b.i – Standard 3, RQ 2

What quadrant is point B in on the coordinate plane?

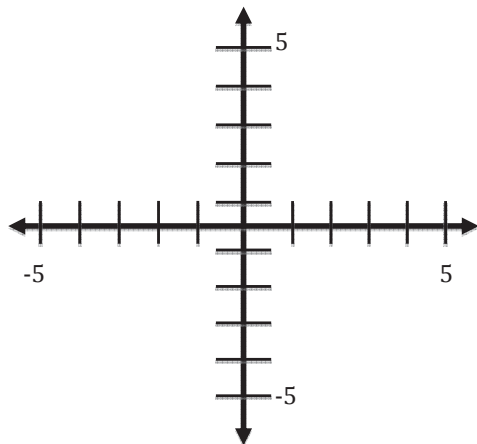


Evaluation Problems

11

114. 6.NS.6.b.i – Standard 3, WQ 1

Plot a point in quadrant 4 on the graph

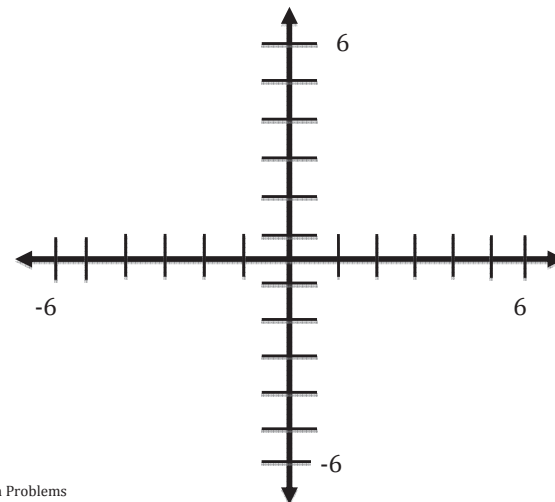


Evaluation Problems

12

115. 6.NS.6.b.i – Standard 3, WQ 2

Plot a point in quadrant 2 on the graph

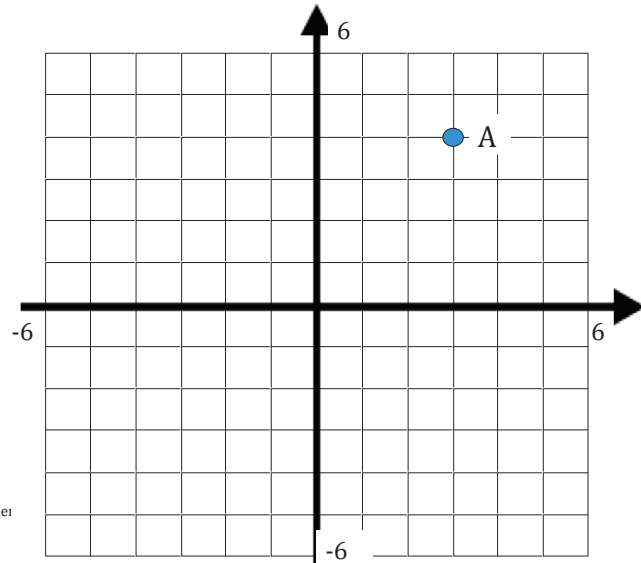


Evaluation Problems

13

116. CCS-6.NS.6.b.ii – Standard 4, RQ 1

What are the coordinates of the reflection of point A over the X axis?

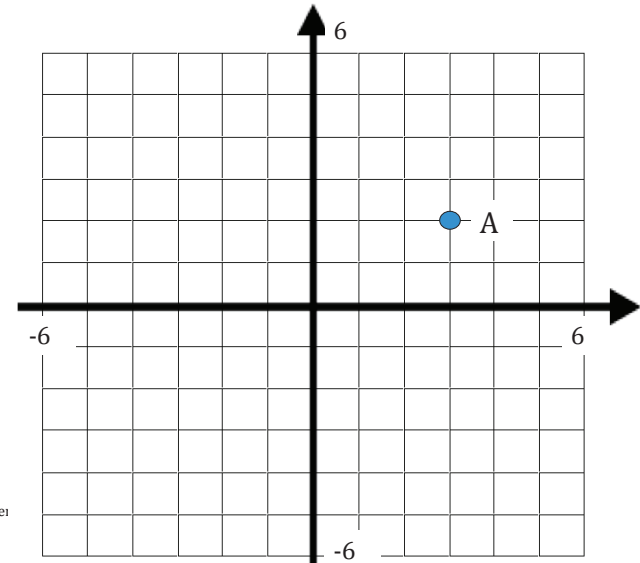


Evaluation Problem

14

117. CCS-6.NS.6.b.ii – Standard 4, RQ 2

What are the coordinates of the reflection of point A over the Y axis?

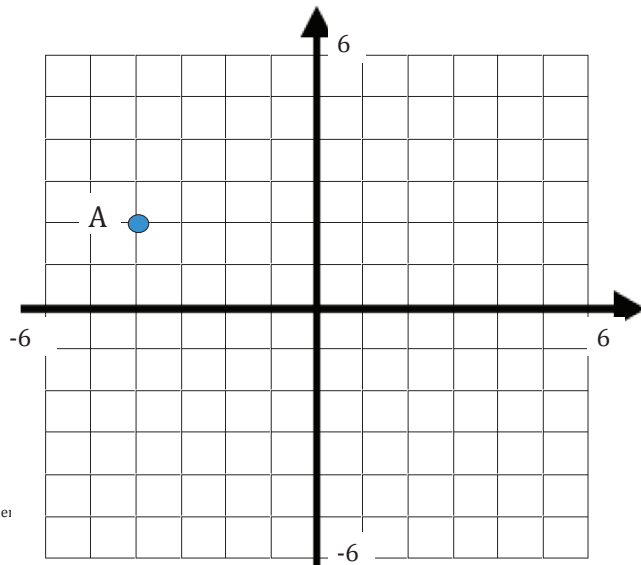


Evaluation Problem

15

82. CCS-6.NS.6.b.ii – Standard 4, WQ 1

Plot the reflection of point A (-4,2) across the X axis on the graph

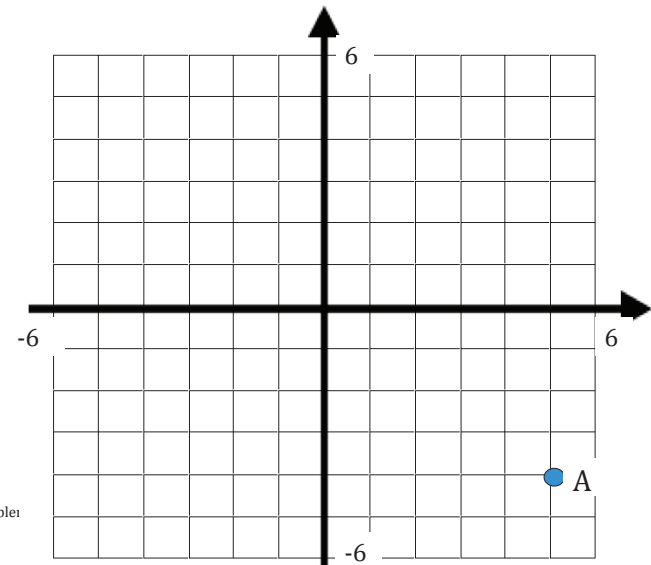


Evaluation Problem

16

83. CCS-6.NS.6.b.ii – Standard 4, WQ 2

Plot the reflection of point A (5,-4) across the Y axis on the graph

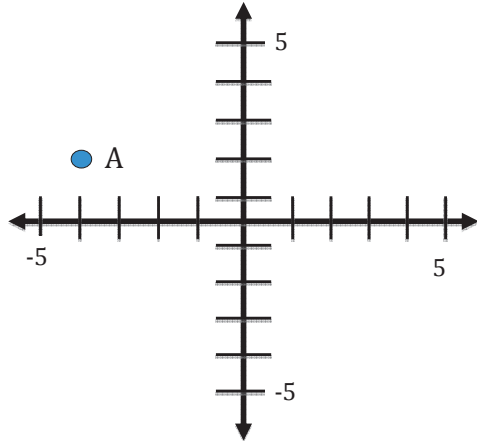


Evaluation Problem

17

105. 6.NS.6.c.ii – Standard 5, RQ 1

What are the coordinates of point A on the graph?

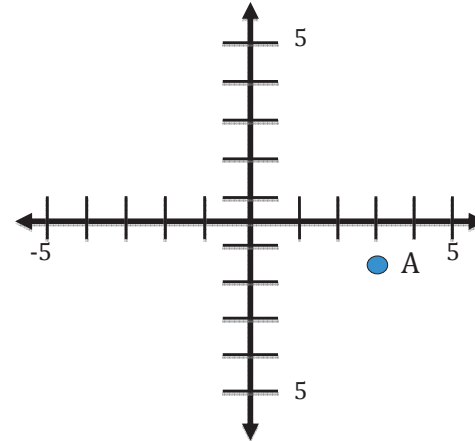


Evaluation Problems

18

106. 6.NS.6.c.ii – Standard 5, RQ 2

What are the coordinates of point A on the graph?

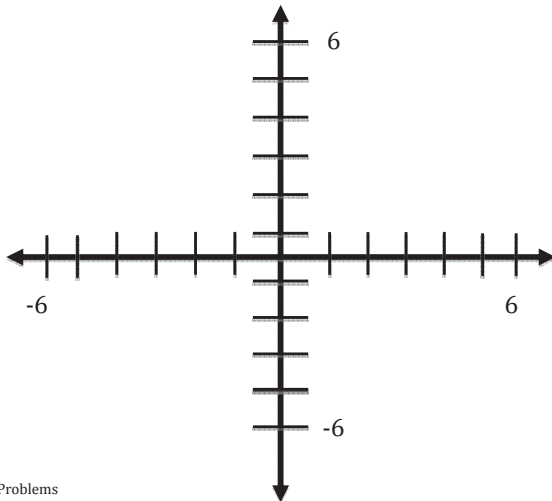


Evaluation Problems

19

108. 6.NS.6.c.ii – Standard 5, WQ 1

Plot (2,-5) on the coordinate plane below.

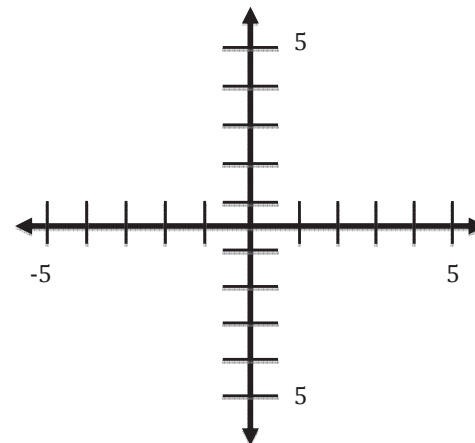


Evaluation Problems

20

109. 6.NS.6.c.ii – Standard 5, WQ 2

Plot (-1,4) on the coordinate plane below.

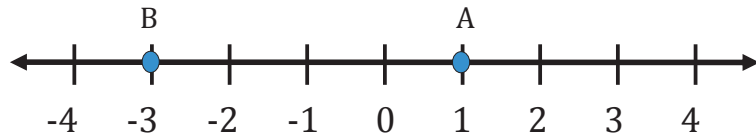


Evaluation Problems

21

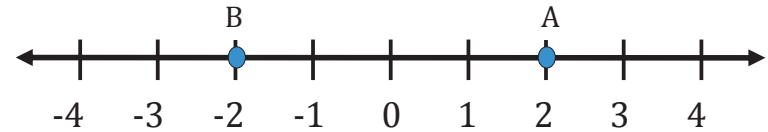
93. 6.NS.7.a – Standard 6, RQ 1

What letter on the number line is the greatest?



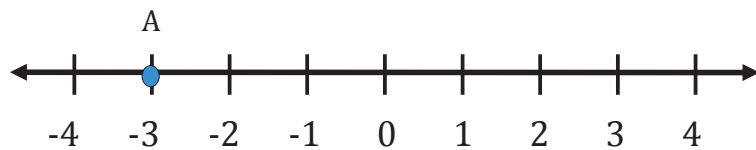
94. 6.NS.7.a – Standard 6, RQ 2

What letter on the number line is the greatest?



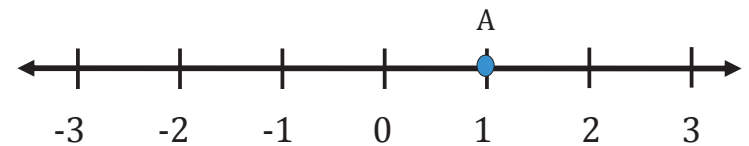
95. 6.NS.7.a – Standard 6, WQ 1

Plot the point that is 4 greater than point A on the number line.



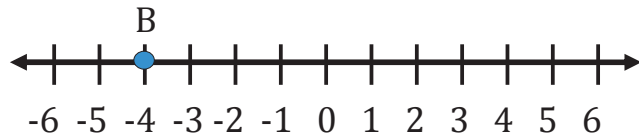
96. 6.NS.7.a – Standard 6, WQ 2

Plot the point that is 2 less than point A on the number line.



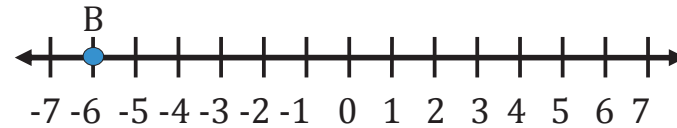
118. CCS-6.NS.7.c.ii – Standard 7, RQ 1

What is the distance from 0 to point B?



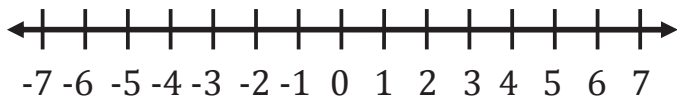
119. CCS-6.NS.7.c.ii – Standard 7, RQ 2

What is the distance from 0 to point B?



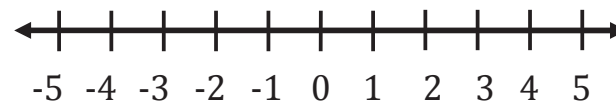
120. CCS-6.NS.7.c.ii – Standard 7, WQ 1

Plot two points at different places on the number line that each have a distance 5 from 0.



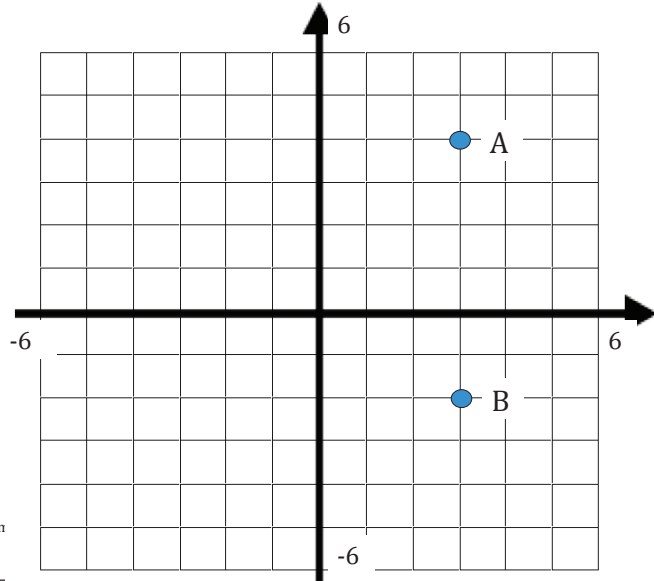
121. CCS-6.NS.7.c.ii – Standard 7, WQ 2

Plot two points at different places on the number line that each have a distance 2 from 0.



122. CCS-6.NS.8.ii – Standard 8, RQ 1

What is the distance between point A and point B on the graph?

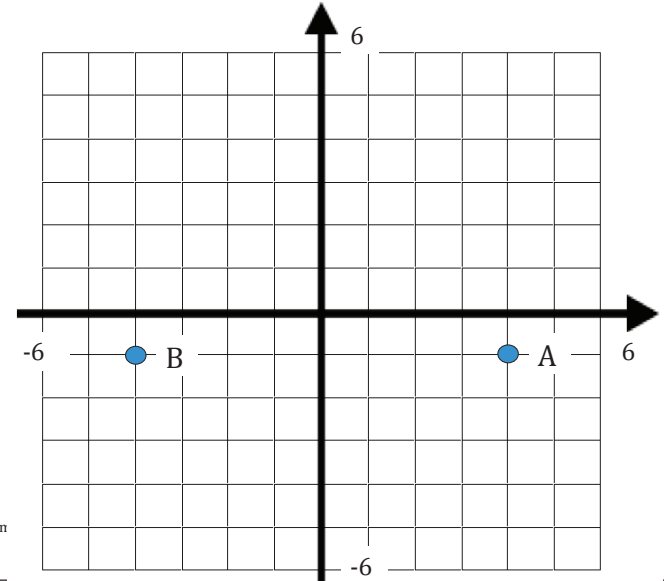


Evaluation Problem

30

123. CCS-6.NS.8.ii – Standard 8, RQ 2

What is the distance between point A and point B on the graph?

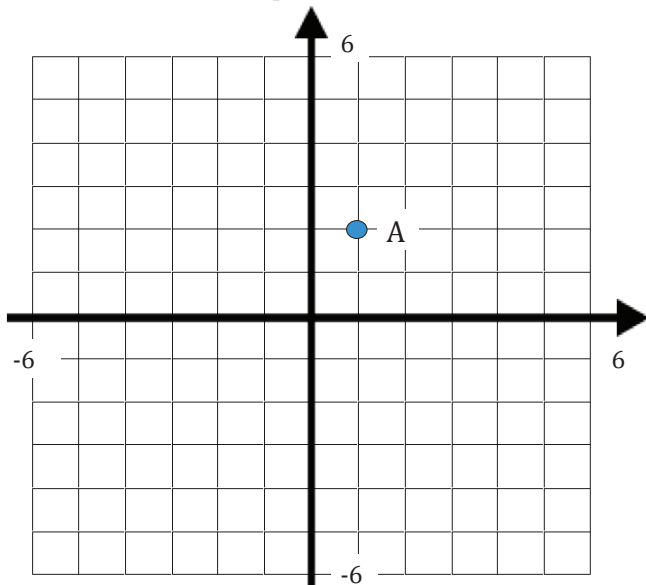


Evaluation Problem

31

124. CCS-6.NS.8.ii – Standard 8, WQ 1

Plot a point 5 units down from point A

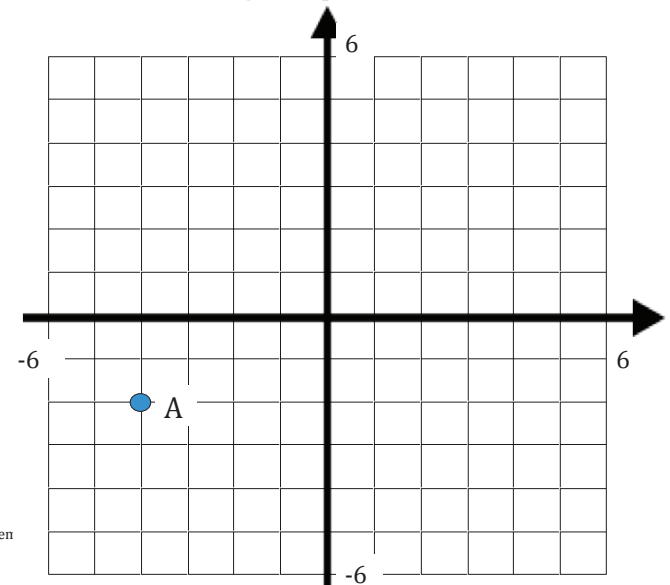


Evaluation Problem

32

125. CCS-6.NS.8.ii Standard 8, WQ 2

Plot a point 6 units to the right of point A



Evaluation Problem

33

5. CCS-6.NS.6A

Plot 6.3 on the number line.

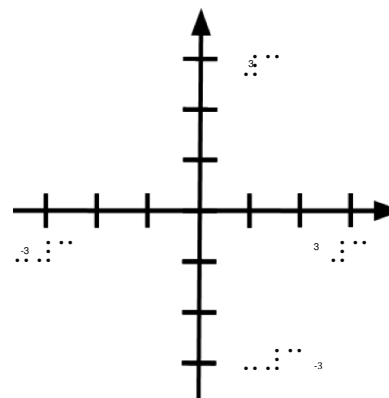
Plot 6.3 on the number line.



7. CCS-6.NS.6B

Plot (-1,2) on the coordinate plane.

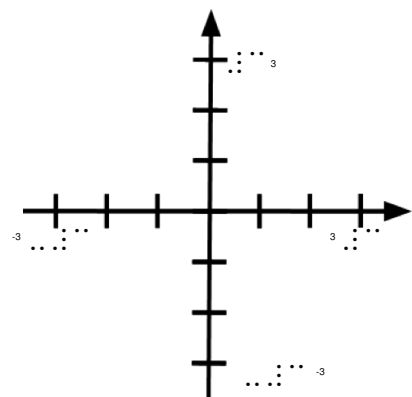
Plot (-1,2) on the coordinate plane.



8. CCS-6.NS.6B

Plot (-3,-2) on the coordinate plane.

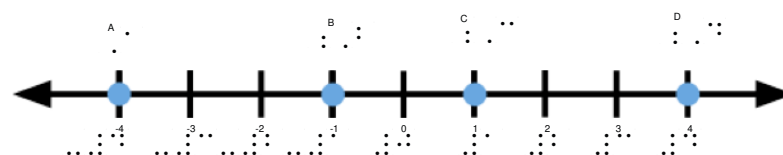
Plot (-3,-2) on the coordinate plane.



9. CCS-6.NS.6.a

What letter is the opposite of:

Q1. A
Q2. B
Q3. C
Q4. D



APPENDIX C

TASK ANALYSIS

This chapter presents the comprehensive results of the task analysis conducted for Section 5.3.

(this appendix is in development and will be released shortly.)

APPENDIX D

SURVEY INSTRUMENTS

This chapter provides examples of the survey and demographics forms used in the studies.

The first document is the survey used in the pilot of Phase 3, Study 1 (classroom situations).

Computers and Graphing Questionnaire (v2)

Experimenter: Please fill out this top portion

Study: _____

Date: _____

Participant Number: _____

Experimenter Name: _____

This survey asks about computer use and graph experience. Please consider a “computer” to be a desktop PC, laptop, or digital note taker. Smartphones and other devices are not “computers” for this survey.

The “mouse” includes a standard mouse or laptop track pad.

General Demographics

1. What is your gender? **Male** **Female**
2. What year were you born? _____
3. What grade are you in school? _____

Impairment

4. Do you have normal hearing or corrected-to-normal hearing? **Yes** or **No**
5. What would you consider to be your current level of visual impairment?
Sighted **Low Vision** **Blind**
6. When did you become low vision? _____
7. When did you become blind? _____
8. Do you have mobility impairments in either arm, including hands?
Neither Arm **Left Arm** **Right Arm** **Both Arms**
9. Which is your dominant hand? **Right** **Left**

Computer Experience

1. How many hours of experience do you have using a computer?

0 1-10 11-100 More than 100

2. About how often do you currently use a desktop computer, laptop, or note taker?

Never Rarely Once a Week Once a Day Many Times a Day

3. Please circle all assistive technologies you regularly depend on to use your computer.

None Magnifier Screen Reader Other

4. Have you used a computer with normal vision or corrected-to-normal vision? If you have always had vision and use a computer, answer “yes”.

Yes No

5. How many hours of experience do you have using a keyboard?

0 1-10 11-100 More than 100

6. How many hours of experience do you have using a mouse?

0 1-10 11-100 More than 100

7. Have you ever used a mouse or laptop track pad? If you are sighted and use a computer, you probably use a mouse. If so, answer “yes”.

Yes No

8. Do you currently use a mouse when using a computer?

Yes No

9. Have you used a mouse when you were sighted? If you are sighted and use a mouse, answer “yes”.

Yes No

10. Please list a few major activities you do with the mouse:

Number Lines and Graphs

11. Number lines are 1-dimensional lines with tick marks and labels, often used in mathematics classes. Have you used number lines before?

Yes No

12. How many hours experience do you have with number lines?

0 1-10 11-100 More than 100

13. What was your level of vision impairment when you learned number lines?

Sighted Low Vision Blind

14. Have you drawn number lines? **Yes No**

15. Coordinate graphs are two-dimensional spaces with tick marks and labels, often used in mathematics classes. Have you used coordinate graphs before?

Yes No

16. How many hours experience do you have with coordinate graphs?

0 1-10 11-100 More than 100

17. What was your level of vision impairment when you learned coordinate graphs?

Sighted Low Vision Blind

18. Have you drawn coordinate graphs? **Yes No**

19. Circle all of the type of graphs that you normally use in class.

Printout Enlarged Printout Tactile Graphics

After School Program

Over the past few weeks, we have explored visual, tactile, and auditory graphs.

- **“Computer graphs”** are the graphs, with sounds, that were completed on the laptops.
- **“Paper/tactile graphs”** are the graphs that were enlarged or made out of tactile graphics.

This section asks for your feedback. Please answer with a scale of 1 to 5, where 1 is strongly disagree, 2 is somewhat disagree, 3 is neutral, 4 is agree, and 5 is strongly agree.

20. ___ In general, number line questions are difficult.
21. ___ In general, graph questions are difficult.
22. ___ I understood the graphing problems with paper/tactile graphs.
23. ___ I understood the graphing problems with computer graphs.
24. ___ I needed the sound to use the computer graphs.
25. ___ I needed my vision to use the computer graphs.
26. ___ I needed my vision to use the paper/tactile graphs.
27. ___ I preferred to be able to hear the sounds with the computer graphs.
28. ___ Between the paper graphs and the computer graphs, I think I was faster on paper.
29. ___ Between the paper/tactile graphs and the computer graphs, I think I had more graph questions correct on paper.
30. ___ Between the paper graphs and the computer graphs, I think I was faster on the computer.
31. ___ Between the paper/tactile graphs and the computer graphs, I think I had more graph questions correct on the computer.
32. ___ I usually had my headphones off.
33. ___ I could often use paper/tactile graphs in my regular math class.
34. ___ I could often use computer graphs in my regular math class.
35. ___ It was easy to understand the meaning of the sounds on the computer graphs.
36. ___ I had to concentrate on the sounds on the computer graphs.
37. ___ Paper/tactile graphs were helpful in solving problems.
38. ___ Computer graphs were helpful in solving problems.
39. ___ Paper/tactile graphs were fun to use.
40. ___ Computer graphs were fun to use.
41. ___ Paper/tactile graphs were annoying to use.
42. ___ Computer graphs were annoying to use.

Computers and Graphing Questionnaire (v2)

43. What are the two most difficult parts of solving number line and graph problems with paper/tactile graph?

1.

2.

44. What are the two easiest parts of solving number line and graph problems paper/tactile graph?

1.

2.

45. What are the two most difficult parts of solving number line and graph problems with computer graph?

1.

2.

46. What are the two easiest parts of solving number line and graph problems computer graph?

1.

2.

47. Do you have any more comments about graphs?

48. Do you have any comments about computer graphs or paper graphs?

49. Do you have any other comments?

Thank you for completing this survey!