Use of the Talking Tactile Tablet in Mathematics Testing

by

Lead authors:

Steven Landau
President of Touch Graphics Company

and

Michael Russell
Senior Research Associate,
Center for the Study of Testing, Evaluation and Educational Policy
Boston College

Co-authors:

Karen Gourgey
Director, Baruch College Computer Center for Visually Impaired People
City University of New York

Jane Erin
Professor, University of Arizona College of Education

Jennifer Cowan
Graduate Assistant, Technology Assessment and Study Collaborative

Resubmitted to Journal of Visual Impairment and Blindness

23 September 2002
Use Of the Talking Tactile Tablet in Mathematics Testing

Abstract:

The article describes an experimental system for administering multiple-choice math tests to students who are blind, visually impaired or otherwise print disabled. Using a new audio-tactile computer peripheral device called the Talking Tactile Tablet, the authors created a preliminary version of a self-voicing test that included twelve items, all of which made reference to a graphical element. Users could take the test, working through the items at their own speed, and learning about associated tactile graphic diagrams by pressing on various features to hear appropriate audio descriptions. A small group of students participated in the system’s evaluation.
Use Of the Talking Tactile Tablet in Mathematics Testing

Over the past decade, testing has become an important component of educational reform efforts. Currently, 49 states have formal testing programs that test students in all public schools on an annual basis. While the subject areas and the grade levels tested vary widely across states, the most recent federal education legislation requires all states to administer annual mathematics and language arts tests to students in grades three through eight.

Recognizing the importance of testing in education, concern about how to test and use scores for students with disabilities has been debated for several decades. As noted by Bennett (1999), the foundation for providing accommodations for students with disabilities was laid by Section 504 of the 1973 Rehabilitation Act which required nondiscrimination on the basis of handicap for all programs receiving federal funding. Four years later, federal regulations stipulated that tests must measure the capabilities of disabled students and not their impairments, with the exception of tests that measured skills that overlapped with the impairment.

While this policy opened the door for providing students with appropriate test accommodations, it also raised questions about the comparability of scores achieved under different conditions and whether scores from students with disabilities should be aggregated with scores for students without disabilities (Thurlow, Scott, & Ysseldyke, 1995a; 1995b). In response, several testing programs began to flag scores for students who received accommodations. In many cases, students with disabilities were systematically excluded from state-testing programs, thus corrupting the aggregate performance of schools and states (McDonnell, McLaughlin, & Morison, 1997; McGrew, Thurlow, Shriner, & Spiegel, 1992; Shriner & Thurlow, 1992; Ysseldyke & Thurlow, 1994). In addition, several states recorded rapid increases in the number of students identified as special education or retained in grade, possibly to avoid participation in state testing programs (Allington & McGill-Franzen, 1992; Haney, 2000; Ysseldyke, Thurlow, McGrew, & Shriner, 1994; Zlatos, 1994).
Concerned that both flagging the results of students with disabilities or outright excluding them from state testing programs may decrease attention, funding and ultimately services provided to students with disabilities, the 1997 Individuals with Disabilities Education Act (IDEA) required that states and districts include students with disabilities in their assessment and accountability programs. As Tindal & Fuchs 1999 (p. 6) summarize, “The assumption is that if schools are to consider the needs of students with disabilities deliberately and proactively in reform and improvement activities, the outcomes of students with disabilities must be represented in public accountability systems.”

For students who are blind or visually impaired, several types of accommodations may be provided, including Braille versions of the test, large print, assistive devices such as a magnifying glass or slate and stylus, or test administrators who read items and responses aloud to the test-taker (Bennett, Rock & Kaplan, 1987; Coleman, 1990; Willingham, Rogosta, Bennett, Braun, Rock & Powers, 1988; Bennett, Rock, & Jirele, 1985; Bennett, Rock, & Kaplan, 1987). Prior research indicates that these accommodations have mixed results on the performance of students and generally finds that accommodations are more useful and valid for language arts tests than for mathematics tests. While little research has focused specifically on mathematics items that require students to work with graphics and/or diagrams, common sense and anecdotal evidence (presented below) suggest that today’s standard accommodations do not provide blind and visually impaired students with adequate access to these graphical elements.

A newly available device, known as the Talking Tactile Tablet (or TTT), holds promise for making graphical elements from multiple-choice math tests more accessible for blind, visually impaired and otherwise print-disabled students (Landau & Gourgey, 2001). The TTT is an inexpensive electronic device that is connected to a host computer via the USB port; a user can mount one of many specially prepared raised-line and textured drawing sheets on the TTT’s touch-sensitive surface, and then can press various shapes, icons and regions on the tactile image to instigate appropriate audio responses. The TTT system relies on a standard arrangement of tools and data entry elements laid
out around a large rectangular work space (see figure 1, a picture of one of the plates created for the math test application). As in Windows-style computing, users become familiar with this design, known as the Tactile Graphic User Interface, and so can quickly and intuitively operate new applications. Currently available applications include a Talking Tactile Atlas of World Maps, a memory/matching game and a curriculum for teaching coordinate geometry. The system has been developed by Touch Graphics Company of Brooklyn, New York (see www.touchgraphics.com), through a series of Small Business Innovation Research grants from the National Institute for Disability and Rehabilitation Research (US Department of Education) and the National Science Foundation.

We created new programming for the TTT for delivery of multiple-choice math tests. Using this system, a student listens to detailed instructions, then scrolls through a
As a test accommodation tool, the TTT offers several advantages. First, it provides students with detailed access to graphic elements and enables them to learn more about specific aspects of the graphics through associated digital voice recordings. Second, the system allows students to work through items independently and at their own pace. Third, students can ask the system to repeat specific information as many times as needed without fatiguing a reader. Finally, since the tactile diagrams describe their features when pressed, the potential audience for this accommodation is larger than for methods that rely on tactile materials annotated in Braille.

But, as Thurlow, McGrew, Tindal, Thompson, Ysseldyke & Elliott (2000) emphasize, before any accommodation is allowed, three conditions must be met. First, it must be established that the accommodation has a positive impact on the performance of students diagnosed with the target disability(s). Second, the accommodation should have no impact on the performance of students that have not been diagnosed with the target disability (that is, providing the accommodation to students with the target disability does
not provide an unfair advantage). And third, the accommodation does not alter the underlying psychometric properties of the measurement scale.

As a first step in exploring the feasibility, utility and validity of using the Talking Tactile Tablet (TTT) as a test accommodation tool for blind and visually impaired students, we undertook a small pilot study that focuses specifically on mathematics items that referenced graphics. As is explained in greater detail below, the study was intended to explore the first and third conditions identified by Thurlow et al. (2000), namely whether the Talking Tactile Tablet has a positive impact on the performance of blind and visually impaired students and how the accommodation impacts the psychometric properties of the test items.

**Study Design**

The pilot study described here examined the extent to which use of the TTT had a positive impact on the performance of students who are blind, visually impaired and/or who have difficulty visualizing graphics and diagrams. To the extent possible, this study also explores the impact the TTT has on item difficulties. The subjects included eight people. As Table 1 summarizes, half of the subjects were able to read Braille, while the other half were not Braille literate and preferred to use large print or a reader. Several of the subjects had also used tactile graphics previously (although not the TTT). Five of the subjects were male and three were female. The subjects’ educational levels ranged from ninth grade to post-college.
Table 1: Description of Participants

<table>
<thead>
<tr>
<th>Participant</th>
<th>Gender</th>
<th>Age</th>
<th>Education Level</th>
<th>Preferred Current Accommodation</th>
<th>Impairment</th>
<th>Prior use of Tactile graphics</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>male</td>
<td>17</td>
<td>11th Grade</td>
<td>Braille</td>
<td>No useful vision</td>
<td>yes</td>
</tr>
<tr>
<td>2</td>
<td>male</td>
<td>17</td>
<td>11th Grade</td>
<td>Large Print</td>
<td>Considerable useful vision</td>
<td>yes</td>
</tr>
<tr>
<td>3</td>
<td>female</td>
<td>14</td>
<td>9th Grade</td>
<td>Braille and reader</td>
<td>no useful vision</td>
<td>yes</td>
</tr>
<tr>
<td>4</td>
<td>male</td>
<td>16</td>
<td>9th Grade</td>
<td>Reader and large print</td>
<td>Considerable useful vision</td>
<td>no</td>
</tr>
<tr>
<td>5</td>
<td>male</td>
<td>19</td>
<td>12th Grade</td>
<td>Reader and large print</td>
<td>Some useful vision</td>
<td>no</td>
</tr>
<tr>
<td>6</td>
<td>male</td>
<td>17</td>
<td>11th Grade</td>
<td>Reader</td>
<td>Considerable useful vision</td>
<td>no</td>
</tr>
<tr>
<td>7</td>
<td>female</td>
<td>24</td>
<td>Bachelor of Science in Math</td>
<td>Braille</td>
<td>No useful vision</td>
<td>yes</td>
</tr>
<tr>
<td>8</td>
<td>female</td>
<td>42</td>
<td>Bachelor of Arts in Math, Master of Science in Rehab Teaching</td>
<td>Braille and reader</td>
<td>Limited useful vision</td>
<td>yes</td>
</tr>
</tbody>
</table>

For this study, three test forms each containing four items were administered to all students. All twelve of the items referenced a diagram or graphical element(s). The items were selected from the publicly released 1998, 1999, 2000, and 2001 Massachusetts Comprehensive Assessment System 10th Grade Mathematics Test. Although the content area categorizations have changed over the years, the items focus on four general areas of mathematics: geometry, measurement, patterns and relations, and statistics and probability (see Table 2).
### Table 2: Summary of Items Included in Form A and Form B

<table>
<thead>
<tr>
<th>Form</th>
<th>Item</th>
<th>Content Area</th>
<th>Short Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>Geometry and Spatial Sense</td>
<td>Measure of Angle in a Pentagon</td>
</tr>
<tr>
<td>A</td>
<td>2</td>
<td>Measurement</td>
<td>Calculate the shaded area of a circle in a square</td>
</tr>
<tr>
<td>A</td>
<td>3</td>
<td>Geometry</td>
<td>Find the most efficient route of a snow plow</td>
</tr>
<tr>
<td>A</td>
<td>4</td>
<td>Patterns, Relations and Functions</td>
<td>Choose graph that best depicts relationship between diameter and circumference of a circle</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>Data Analysis, Statistics and Probability</td>
<td>Graphical interpretation of a car trip</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>Geometry</td>
<td>Find the length of a line in a triangle</td>
</tr>
<tr>
<td>B</td>
<td>3</td>
<td>Geometry</td>
<td>Kite flying – calculate the length of one side of a triangle</td>
</tr>
<tr>
<td>B</td>
<td>4</td>
<td>Data Analysis, Statistics and Probability</td>
<td>Determine the probability of a dot landing in a shaded area of a circle</td>
</tr>
</tbody>
</table>

Two of the test forms (referred to as Form A and Form B in Table 2 and throughout the remainder of the text) were administered both in the standard accommodation and on the TTT. The third form, which was used for practice, contained items which we considered “experimental” and was administered only on the TTT. These items were experimental in that two items required students to use a numerical key pad to supply a response rather than to select a response (that is, it was an open-ended rather than multiple-choice item). A third item required students to reference multiple diagrams and then press on the diagram rather than use the menu to select their response. The fourth item was a standard multiple choice item. These three experimental items were included on this form in order to explore the feasibility of using the TTT for open-ended responses requiring students to supply a numerical response and for using the
diagram to select responses. Since these items were considered experimental, performance results on the practice form were not analyzed. In addition, the practice form was administered before students performed either Form A or Form B on the TTT, and thus provided an opportunity for students to become accustomed to working with the TTT.

As suggested by Thurlow, et al (2000), a two-group design was employed. As Table 3 indicates, the four Braille and the four non-Braille readers were randomly assigned to one of two groups. Students in Group A were provided with the current accommodations – a choice of using a Braille version, having the items read aloud, and/or using an assistive device that magnified text and graphics – for Form A. For Form B, students in Group A used the TTT. Conversely, Group B used the TTT for Form A, and the preferred current accommodation for Form B. Thus, this experiment compared the effect of using the TTT to the preferred current method of accommodating visually impaired students.

Table 3: Pilot Study Research Design

<table>
<thead>
<tr>
<th></th>
<th>Group 1</th>
<th>Group 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Current Accommodations</td>
<td>Form A</td>
<td>Form B</td>
</tr>
<tr>
<td>With TTT</td>
<td>Practice Test</td>
<td>Practice Test</td>
</tr>
<tr>
<td>With TTT</td>
<td>Form B</td>
<td>Form A</td>
</tr>
</tbody>
</table>

All students first completed the prescribed form (A or B) using their preferred current accommodation. Students then completed the Practice Test on the TTT before taking the remaining form on the TTT. Given the total time of testing and the effort required to learn how to use the TTT while working on the Practice Test, there is a possibility that fatigue affected performance on the final form administered on the TTT.
The researchers recognize that a better experimental design would have randomly divided both groups in half, with one sub-group taking Test Form A first and the second sub-group taking Test Form B first.

In addition to comparing the performance of students when provided with their standard accommodation and when using the Talking Tactile Tablet, three other types of data were collected. First, prior to finalizing the TTT versions of the mathematics tests, a focus group was held with six blind and visually impaired adults, one of whom later participated in the TTT experiment. The purposes of this focus group were twofold. First, we hoped to develop a better understanding of the types of problems blind and visually impaired students encounter when taking mathematics tests in Braille, with a reader and/or with other devices. Second, we hoped to acquire preliminary feedback on the design of the TTT and features that should or should not be included in the delivery system.

Second, students were videotaped as they performed the mathematics items using their standard accommodation and when using the TTT. The video tapes were used to identify problems students appeared to have accessing the mathematics items (specifically the graphical elements) as well as technical problems that arose as students used the TTT interface to navigate through the test.

Third, students were interviewed after completing all forms. The purposes of these interviews was to document problems they encountered while working on the mathematics items while using the TTT and when offered their choice of the current accommodations, their preference after having used the TTT, and suggestions on how to improve the TTT.
Findings

In this section we present the findings from each data source. A more general discussion of the results is then presented.

Focus Group

During the focus group, all participants indicated that they had experienced problems when taking mathematics tests. Those who preferred to use Braille versions of the tests indicated that it was often difficult to read graphics presented in tactile form. Some participants also described the need for proctors to re-draw the diagrams using tactile tools so that they could better understand the diagrams. One participant (who has a degree in mathematics) indicated that some items on the SAT and GRE were not administered to her because it was too difficult to create tactile versions of the accompanying graphics.

Other participants who had used or who had seen students use readers indicated that several problems arose. These problems fell into three categories. First, participants indicated that the quality of the readers varied widely. In some cases, the readers did not speak English as a first language. As a result, ESL readers occasionally mispronounced words. In other cases, the readers were not familiar with the test content or terms and would also mispronounced or mis-read words.

Second, participants indicated that readers sometimes gave hints to the correct answer. At times these hints were provided intentionally with readers making comments such as, “Are you sure” or “You might want to check that again.” At other times, the hints were more subtle and probably unintentional with the reader pronouncing an option or key word with more emphasis or with a different tone to their voice.

Third, participants indicated that they and/or their students were sometimes reluctant to ask proctors to re-read parts of an item when they were not sure what was being asked or what a response was. This seemed to be more of an issue for language arts items that required examinees to read extended passages, but participants noted that the problem also occurred on mathematics tests.
Several participants expressed frustration about the difficulty they and/or their students have in accessing diagrams and other graphical elements. One participant indicated that she had spoken with officials in the state department of education (NY), but that “they seemed largely unaware of the issues that visually impaired people encounter taking the tests.”

One participant indicated that many of the problems that others identified were not a problem of technology but of translation. This participant was unsure whether a computer-based system would result in better translations (from printed text to Braille or to another form that is accessible to blind and visually impaired students). This participant, however, was in the clear minority as all other participants believed that the TTT held promise for improving access.

Braille readers, however, expressed concern that the current version of the TTT tactile plates do not include labels presented in Braille but instead required users to press on an object which the computer then verbally identified. These participants felt that they could develop an understanding of the diagram more quickly by reading labels in Braille. Others, including Steven Landau, the designer of the TTT, expressed concern that if Braille labels were included, it might clutter the diagram and/or cause confusion for examinees that could not read Braille. In the final version of the test materials, limited Braille tagging was added to the tactile diagrams, but to get a full understanding, it was necessary to make use of the “touch and tell” feature.

The final concern raised by several participants was the feasibility of using a device such as the TTT only for testing. Participants felt strongly that nothing new should be introduced during the test. The participants strongly urged the study to provide students an opportunity to use the TTT for several days prior to the actual testing so that they were comfortable and knew how to use the TTT, particularly for math. The designer of the TTT, however, informed them that besides the test items themselves, no other mathematics curriculum had yet been developed to work with the TTT. For this reason it would not be possible for students to use the TTT for mathematics prior to the study.
Nonetheless, the participants made a valid point, namely that the validity of scores may be negatively affected due to students lack of comfort with the TTT.

Test Results

Figure 2 summarizes the results for each item and test form and compares these to the item difficulty (percent of students succeeding on the item) as reported by the Massachusetts Department of Education. Among the eight items, students performed better on five items when using the TTT and performed the same on the remaining three items. On both test forms, participants performed better when they used the TTT as compared to their preferred current accommodation.

**Figure 2:** Percent of students succeeding on items by accommodation

![Performance on tenth grade math MCAS items](image)
It is also interesting to note that using the TTT yields item difficulties that more closely resemble the item difficulties obtained during the actual MCAS administrations. As Figures 2 shows, the percentage of students succeeding on each item are noticeably lower than the state averages when examinees used the current accommodation. However, the item difficulties for six of the eight items administered on the TTT are close to those reported by the state.

To summarize, when examinees used the TTT, the results indicate that students tended to perform better and that the resulting item difficulties more closely resembled those reported for the actual MCAS administration.

**Video Observations**

As described above, students were video taped as they performed each test form. These videos were reviewed to examine problems that students seemed to encounter while performing the test forms.

In general, it did not appear that students had any major problems while taking the tests using the current accommodations. At times, students asked for clarification as to what a question was asking or what a diagram depicted. But after input from a proctor, the students were able to work on the problem and provide a response.

When working with the TTT, a handful of technical problems occurred. During both the Practice Test and the actual test, the program jumped from an item to the main menu without the student using the menu buttons. When this occurred, students had to use the navigation buttons to return to the item and then resume working.

With the exception of two examinees, students were able to place the tactile test forms onto the tablet, calibrate the tablet and then press the correct tablet form identification keys. Repeatedly we were surprised by how quickly examinees mastered the calibration process and were able to calibrate the second form with ease. In one instance, however, the TTT seemed to allow the student to proceed without completing the calibration process. It was unclear whether the TTT simply did not play the accompanying sound that indicates that the proper calibration button was pressed or
whether the TTT advanced without the student actually pressing the button. In either case, the student appeared confused by the result and the test administrator had to restart the calibration process. A second student experienced problems calibrating the TTT when they pressed the wrong button but the TTT made sounds as if it were properly calibrated. Again, the test administrator intervened and re-started the calibration process.

Finally, several students experienced trouble with the experimental items. Specifically, a few students were unsure how they were to enter responses for the open-ended items and were only able to do so after clarification from the test administrator. In addition, students were confused how to respond to the experimental item that required them to press on the correct diagram as opposed to using the menu to select an option.

Despite these technical problems (all of which should be easy to correct and could be identified in advance through more extensive beta-testing), we cannot emphasize strongly enough how quickly seven of the eight students were able to learn how to interact with the TTT. Although none of the students had prior experience with the TTT, after working on the practice test, seven of the eight students seemed comfortable and facile with the TTT while working on the actual tests.

**Student Interviews**

During the interviews, students identified several aspects of the system that they would like to have improved or customized. As noted above, several students experienced trouble with the system jumping from an item to the main menu. Although they were able to return to the item, they noted that it would be easier to work on an item without being interrupted.

One student also noted that it would have been easier to work on problems if the system did not periodically repeat the question or prompt her while s/he was thinking. This student wanted to hear the question and option, then work on her/his solution in silence, and then re-activate the voice to assist her in selecting her response.
A few students also wanted greater control over the speed with which the system read questions and answers. At times, the students felt the voice could have been faster (e.g., during the calibration process). But when reading the question and options, some students wanted the voice to proceed more slowly.

One examinee, who was a Braille reader, indicated that it would have been helpful to have responses and various parts of the tactile surface labeled in Braille rather than being read. Similarly, the student who used the CCTV to magnify the paper form of the items wanted the graphics to be larger on the TTT version and for colors to be used to differentiate different aspects of the graphics.

Finally, after using the TTT, only one student indicated that s/he would still prefer to take tests using the current accommodations.

Summary

This pilot study was undertaken to explore the feasibility and impact of using the Talking Tactile Tablet as a test accommodation for blind and visually impaired students. In evaluating our results, it is important to remember that this was only a pilot study that focused on a small sample of students.

As described above, the TTT had a positive impact on student performance. On the majority of items, students performed better when using the TTT as compared to current accommodations. On three items, the performance difference was very large. On both forms, the difference in performance was of practical significance.

Although the small sample sizes limited our analysis of the impact the test accommodations had on the psychometric properties of the items, it does appear that the item difficulties for items administered on the TTT more closely resemble those reported on the actual MCAS.

It should be noted that these differences in performance occurred despite the potential impact of fatigue and the technical errors that occurred. In addition, although
most students quickly learned how to use the system, the TTT was new for all students. In future studies, it is strongly suggested that students be allowed to work with the TTT prior to testing and that more thorough beta-testing be performed prior to testing.

In summary, we believe that the results reported here indicate that the TTT holds promise as a test accommodation tool for blind and visually impaired students. For this reason, we strongly recommend that additional research be conducted that includes a larger sample of students and addresses the second criteria of Thurlow et al (2000), namely that the accommodation should not have an impact on the performance of students that have not been diagnosed with the target disability. This can be accomplished by testing the TTT on a group of sighted students.

References


