Multimodal feedback: an assessment of performance and mental workload

H. S. Vitense†, J. A. Jacko‡ and V. K. Emery‡

†Medtronic, 7000 Central Avenue N.E., B173, Minneapolis, MN 55432, USA
‡School of Industrial and Systems Engineering, Georgia Institute of Technology, 765 Ferst Drive, Atlanta, GA 30332-0205, USA

Keywords: Multimodal; Auditory; Haptic; Visual; Feedback; Human–computer interaction.

Multimodal interfaces offer great potential to humanize interactions with computers by employing a multitude of perceptual channels. This paper reports on a novel multimodal interface using auditory, haptic and visual feedback in a direct manipulation task to establish new recommendations for multimodal feedback, in particular uni-, bi- and trimodal feedback. A close examination of combinations of uni-, bi- and trimodal feedback is necessary to determine which enhances performance without increasing workload. Thirty-two participants were asked to complete a task consisting of a series of ‘drag-and-drops’ while the type of feedback was manipulated. Each participant was exposed to three unimodal feedback conditions, three bimodal feedback conditions and one trimodal feedback condition that used auditory, visual and haptic feedback alone, and in combination. Performance under the different conditions was assessed with measures of trial completion time, target highlight time and a self-reported workload assessment captured by the NASA Task Load Index (NASA-TLX). The findings suggest that certain types of bimodal feedback can enhance performance while lowering self-perceived mental demand.

1. Introduction

Effective and efficient communication between a human and a computer is the critical link driving the usefulness of high-performance systems. However, performance enhancements cannot be achieved in the presence of excessive user workload. The research presented here is an investigation into methods designed to broaden and fortify the links of communication without increasing user workload in the context of multimodal interfaces.

Multimodal interfaces are one way to broaden and strengthen the critical communication between humans and computers. These interfaces have the potential to enhance the user’s interaction with information technologies through an augmentation of the perceptual processes via multiple modalities. Owing to the growing complexity of technology and applications, in some situations a single modality no longer allows users to interact effectively across all tasks and environments (Oviatt et al. 2002). Consequently, multimodal input is often preferred over unimodal input (Oviatt et al. 1997). These interfaces make the most effective use of the variety of human sensory channels, alone and in combination. In addition to input, output can also be multimodal in nature. Often multimodal output is referred to as multimodal feedback.
This paper describes research aimed at investigating how multimodal feedback affects user performance. The literature documents the need for a better understanding of interfaces that use more than two feedback modalities. Consequently, this research seeks to establish new recommendations for bi- and trimodal feedback. A novel multimodal interface was used that employed auditory, haptic and visual feedback. By understanding how these forms of feedback affect users’ performance, general recommendations are provided that increase the effectiveness and efficiency of direct manipulation tasks.

2. Background

2.1. Multimodal interfaces

Bolt’s ‘Put That There’ system is often cited as the first multimodal interface (Bolt 1980). Bimodal in nature, Put That There integrated speech and hand gestures (i.e. pointing on a touchpad) to create and move objects on a two-dimensional display (Bolt 1980). Table 1 is a concise chronological summary of examples of multimodal interface research, commencing with Bolt’s bimodal interface. As speech recognition technology matured during the 1980s and 1990s, speech input coupled with keyboard or mouse input was widely investigated. An example can be seen in Cohen

<table>
<thead>
<tr>
<th>Multimodal Input Research</th>
<th>Multimodal Output Research</th>
</tr>
</thead>
<tbody>
<tr>
<td>System</td>
<td>System</td>
</tr>
<tr>
<td>Speech recognition and gesture</td>
<td>Speech synthesis and haptic</td>
</tr>
<tr>
<td>Speech recognition and mouse</td>
<td>Non-speech auditory and haptic</td>
</tr>
<tr>
<td>Speech recognition and head movement</td>
<td>Non-speech auditory and visual</td>
</tr>
<tr>
<td>Smith et al. (1996)</td>
<td>Brewster (1998a,b), Brewster et al. (1994),</td>
</tr>
<tr>
<td>Speech recognition and pen</td>
<td>Brewster and Crease (1999)</td>
</tr>
<tr>
<td>Speech recognition with pen and mouse direct manipulation</td>
<td>Akamatsu (1992), Arsenault and Ware (2000), Campbell et al. (1999), McGee (1999), Oakley et al. (2001)</td>
</tr>
<tr>
<td>Speech recognition and lip reading</td>
<td></td>
</tr>
<tr>
<td>Ren et al. (2000)</td>
<td></td>
</tr>
<tr>
<td>Head movement and gesture</td>
<td>Ma et al. (2000)</td>
</tr>
<tr>
<td>Keates et al. (2001)</td>
<td></td>
</tr>
</tbody>
</table>
et al. (1989) and their work with Shoptalk, which incorporated natural language and direct manipulation. In part due to the advancement of technology, great strides have been made since the realization of these initial multimodal interfaces.

As previously mentioned, multimodal interaction can take the form of input into the computer and/or output (i.e. feedback) from the computer. Thus, multimodal research focuses either on input or output. Table 1 summarizes the different levels of attention given to input research versus output research over the past 20 years. In fact, the majority of work in this field to date has concentrated on how users performed with each of the input modalities separately. Input modalities have, at times, been observed in combination to determine optimal configurations for particular tasks. Research and development in multimodal output has also been limited and primarily focused on investigations involving bimodal conditions.

An examination of existing HCI research in the multimodal literature, such as shown in table 1, reveals a distinct gap in the knowledge base with respect to research and development of multimodal output. It is essential that multimodal feedback not be under emphasized in the continued development and advancement of multimodal interfaces.

2.2. Auditory, haptic and visual feedback research
Auditory, haptic and visual feedback represent the most prominent types of feedback. These modalities were selected for this study due to their current and potential use in contemporary Graphical User Interfaces (GUIs). Of the three types, visual feedback has been most commonly employed. Auditory and haptic modalities were used in this study because of their tremendous potential for contributing to effective interaction when the visual modality is compromised, engaged, or overwhelmed. Sound has long been used to focus users’ attention (Sanders and McCormick 1993). The use of haptic feedback, however, is relatively new, especially with respect to maintaining effective human–computer interaction. The most extensive use of haptic feedback has been within virtual reality domains (e.g. Akamatsu and MacKenzie 1996, Liu et al. 2000).

2.2.1. Auditory feedback: In this research, auditory feedback was provided with an earcon designed to indicate movement (Blattner et al. 1989). Earcons are abstract, musical tones that represent parts of an interface and have been used to provide cues for scanning and conveying navigational cues (Brewster et al. 1996, Brewster 1998b). They have also been shown to enhance performance while performing typical functions within a GUI such as using a scrollbar (Brewster et al. 1994).

2.2.2. Haptic feedback: Haptic feedback was conveyed in this study via vibration provided by a haptic mouse. The earliest work with haptic feedback dates for the 1990s. In 1994, Akamatsu and Sato developed a haptic mouse that produced ‘tactile feedback’. At the same time, Engle et al. (1994) found that directional two degrees of freedom force feedback from a trackball improved speed and error rates in a targeting task. Since then, not only has the literature-base grown, but also has the development and production of off-the-shelf haptic devices. There are a number of off-the-shelf haptic devices available for personal computers.

2.2.3. Visual feedback: Finally, visual feedback manifested itself in this study in the form of blue highlighting. The history of GUIs has relied on visual feedback of various forms. Visual feedback provides location and identification information as
well as cues that signify correct actions. Because of its importance, multimodal research has heavily investigated the effects of adding supplementary modalities to the visual modality.

2.2.4. *Supplementing visual feedback*: Brewster and colleagues have conducted considerable research on the effects of non-speech audition (i.e. earcons) within GUIs. In 1994, Brewster *et al.* tested an auditory-enhanced scrollbar versus a typical visual scrollbar. They found that by adding earcons to the function of the scrollbar, the time to complete the task and mental effort exerted were both reduced. Brewster *et al.* (1996) found that earcons could successfully supplement visual feedback in a scanning task. Sonically enhanced drag-and-drop tasks also yielded better performance than traditional visually enhanced drag-and-drop tasks (Brewster 1998a). These studies show that the usability of GUIs is improved by the addition of auditory feedback to the already present visual feedback.

Visual feedback has also been augmented with haptic feedback. Typical GUI tasks such as moving a scrollbar, selecting an item from a menu, and pointing and clicking were tested with visual and haptic feedback (e.g. McGee 1999, Keates *et al.* 2001, Oakley *et al.* 2001). These studies underline the need for further study of haptic feedback within multimodal interfaces. They have also illustrated that additional modes of feedback, when added to visual feedback, benefit user performance for some tasks.

2.3. *Performance measurements*

The study of human–computer performance has evolved over decades (Nickerson and Landauer 1997). Designers, researchers and practitioners have always struggled to report quantifiably how well a system performs. There are two ideas traditionally related to performance: effectiveness and efficiency. Effectiveness can be thought of as ‘doing the right things’, whereas efficiency is ‘doing things the right way’ (Salvendy and Carayon 1997). It is vital to investigate both effectiveness and efficiency because they are inherently related. The study of human performance addresses how humans perceive information, what they decide to do with it, and how actions are carried out (Wickens 1992). To assess human performance adequately in this study, it was necessary to capture both mental workload and measurements of time.

2.3.1. *Mental workload measurement*: Mental workload has long been recognized as an important element of human performance in complex systems (Gopher and Donchin 1986). Moray (1988) found that the optimization of mental workload could reduce human error, improve system safety, increase productivity and increase operator satisfaction (as cited in Xie and Salvendy 2000). Mental workload has a direct influence over a user’s ability to perform tasks and can therefore impact the effectiveness and efficiency of interactions with computers.

Owing to its high validity, applicability and popularity in similar usability research, NASA-TLX was used in this research to assess subjective workload (Brewster *et al.* 1994, Pitt and Edwards 1997, Miller and Hannen 1999). Byers *et al.* (1988) investigated NASA-TLX, SWAT, Overall Workload and the Modified Cooper-Harper scale in an effort to measure each technique’s validity. NASA-TLX was found in this study, as well as in others, to be the most valid measure of subjective workload, to have the highest user acceptance, and to have the smallest
between-subject variability (Vidulich and Tsang 1986, Byers et al. 1988, Hart and Wickens 1990, Hill et al. 1992). NASA-TLX ratings have also been shown to be sensitive to experimentally manipulated levels of workload and be more reliable than other subjective techniques (Aretz et al. 1996). The NASA-TLX is a multidimensional, self-reported assessment technique that provides an estimation of the overall workload associated with task performance. The assessment is comprised of the relative contribution of six underlying psychological factors to overall workload including mental demand, physical demand, temporal demand, performance, effort and frustration level (Hart and Staveland 1988). The respondents provided absolute ratings on each subscale immediately after completing a designated task. In this study, mental demand was of primary interest and thus is the only factor reported here. This subjective assessment of workload provides a useful system measure when combined with objective measures such as performance times.

2.3.2. **Time measurements**: Trial or task completion time has frequently been used to gauge differences between unique modalities and combinations of modalities applied in multimodal interfaces. For example, Brewster’s (1998a) research used total task completion time to identify differences between auditory feedback and visual feedback during a drag-and-drop task. Rosenberg and Brave (1996) used trial performance time to quantify the effect of a force feedback joystick. Oakley et al. (2001) also used task completion time to examined users’ reactions to haptic and visual feedback as they selected items from a menu in a GUI. These examples of trial completion time represent the general, overall, performance time required to complete one trial. Trial completion time was captured in this study. A second, more specific, measurement of performance also captured in this study was target highlight time. In Brewster (1998a), target highlight time was measured for each drag-and-drop. Target highlight time represented the time the target and source icons were overlaid before the user dropped the source icon. In other words, target highlight time captured the time that began when a user positioned the source file icon over the target folder, which triggered feedback, and ended when the source file was released (i.e. mouse button released). Target highlight time is a more specific measure of performance because it captured a subset of the trial completion time; that portion that is most influenced by the feedback itself. Brewster also paired these time measures (i.e. trial completion time and target highlight time) with the NASA TLX workload index to provide a more complete indication of how the different modalities affected the HCI. This present study is an extension of Brewster (1998a), with specific attention given to the impact of the inclusion of a third modality, haptic feedback.

3. **Methods**

3.1. **Participants**

Thirty-two participants were selected for inclusion in this study. Their ages ranged from 21 to 36 years. The participants were monetarily compensated for their participation. The inclusion criteria were threefold: (1) right-handedness: to be able to use the right handed haptic and ordinary mouse provided; (2) corrected vision near 20/20: to be able to see the visual feedback; and 3) near normal hearing capability: to be able to detect the auditory earcon generated by the computer.
3.2. Apparatus
The Auditory–Haptic–Visual (AHV) Multimodal software, a customized product developed for this project, facilitated the drag-and-drop task and the manipulation of the three modalities within a GUI. The AHV Multimodal Software was run on a Pentium IV 2.0-GHz computer with a 21-inch colour monitor set at a resolution of 1152 × 864. The participants were positioned an average 2.5 feet from the monitor. A Logitech WingMan® Force Feedback Mouse was used to provide the haptic feedback. The haptic feedback manifested itself in the form of vibration. During feedback conditions in which the software did not, by design, provide haptic feedback, the participants were instructed to use the ordinary Microsoft® mouse. A simple calibration test revealed no statistical differences between the operation of the two mice for each participant.

3.3. Procedure
The participants were given an explanation of the research and their role in the study from the experimenter. Each individual was screened to meet the visual, auditory and handedness criteria. Upon satisfaction of the criteria, the participants were given instructions on how to complete the NASA-TLX subjective workload ranking form. Finally, the participants received task-specific familiarization training.

3.3.1. Training: The goal of the training task was to familiarize the participants with the experimental task and each feedback condition so that by completion of the training the participants were equally proficient with each feedback condition. Because auditory and haptic feedback are less common than visual feedback, there was a potential for biased performance in the conditions containing visual feedback. This potential for bias was eliminated by the training task. The statistical analyses run on the training task data confirmed none of the feedback conditions were statistically different with respect to participants’ performance. Consequently, the participants began the experimental task equally proficient in the auditory, haptic and visual feedback conditions. Once the participants completed the training task and felt comfortable with each feedback condition, they were given an explanation of the actual experimental task. Each participant was presented with a notification screen to inform them about what type(s) of feedback they should anticipate before every upcoming condition.

3.3.2. Experimental task: The experimental task evaluated in this study was a direct manipulation task similar to that used in Brewster (1998a), consisting of a search selection drag-and-drop operation. A 5 × 5 grid of Microsoft Word® file icons, 25 in total, appeared in the middle and upper portion of the screen while five Microsoft Office® folder target icons appeared at the bottom of the screen (figure 1). The file and folder icons were lettered A–E. The participants were asked to move each lettered file icon to the folder icon with the same letter (A–E) as ‘quickly and accurately’ as possible. The locations of source and target icons were manipulated throughout the experimental trials. The conditions for which the feedback was manipulated, varied randomly.

Upon placing a file in a folder, the lettered file icon disappeared. At that same time, a replacement file reappeared in the file’s original position to serve as a placeholder. This was done to ensure search complexity remained consistent for the duration of the trial. The replacement files were labelled ‘File X’ and were immobile.
to avoid confusion as to which files had yet to be moved into folders. Once all the files were placed in folders, the screen was cleared and was replaced with a new arrangement of file and folder icons. The participants completed two replications of the task within each feedback condition, and then completed a NASA-TLX assessment.

3.4. Experimental design

This study was a $2 	imes 2 	imes 2$ factorial, within-subject design. The independent variables were auditory feedback, haptic feedback and visual feedback. Mental workload, trial completion time and target highlight time were the dependent variables.

Each of the three independent variables had two levels: ‘present’ and ‘absent’. An independent variable with the level ‘present’ indicates that the associated feedback was provided in the condition, and a variable with a level ‘absent’ means the associated feedback was not provided. Table 2 shows the different variables and levels that comprise each of the eight conditions. A Latin square design was used to counterbalance the various sequences in which the levels of the independent variables occurred (Keppel 1991). This design was repeated four times to accommodate 32 participants.

In one of the conditions, no feedback was provided. In three conditions, unimodal feedback (i.e. visual, haptic and auditory) was tested. In another three conditions, bimodal feedback (i.e. haptic/visual, auditory/visual and auditory/haptic) was tested. Lastly, a trimodal feedback condition (i.e. auditory/haptic/visual) was tested.
3.4.1. **Independent variables**: The auditory feedback conditions were generated by an earcon (Brewster 1998a). When the participant positioned the file icon over the folder target icon, a musical tone sounded. This sound alerted the participants to the fact that they were indeed over the target. So as not to complicate the audio feedback, the same tone, representing ‘move’, was used for each target.

A vibration effect from the force feedback mouse provided the haptic feedback. The vibration effect can best be described as slight, quick movements. When the file icon was over the folder icon, the force feedback mouse vibrated. The vibration was generated by a triangle waveform for 0.30 s.

The visual feedback condition was in the form of a blue highlight. When the file icon was over the desired target, the target icon highlighted blue. This visual feedback, like the other types of feedback, signified that the participant had correctly moved the file to that particular folder. These three types of feedback, auditory, haptic and visual, were presented alone and in combination.

3.4.2. **Dependent variables**: Performance was assessed with three measures: mental workload, trial completion time (TCT), and target highlight time (THT). The NASA-TLX workload measure, administered after each feedback condition, measured mental workload qualitatively. The mental demand subscale gauged the participants’ assessment of the underlying psychological indicators of workload associated with the experimental task, and more specifically the ability to perform the drag-and-drop task. After reviewing the rating scale definitions, participants gave an assessment by providing an absolute rating on a scale from 0 (no workload) to 100 (extremely high workload).

The TCT was the time it took a participant to complete one trial within the experimental task. The second variable collected to measure performance time was THT, which was the time between when a participant positioned the file over the folder and when they released the mouse button, thus dropping the file into the folder. THT was collected every time a file was dropped into a folder. Both performance times, reported in seconds, were captured by the computer running the AHV Multimodal software. Statistical analyses of these measures revealed how user performance varied with respect to the type of feedback presented.

### 4. Results

A general linear model repeated-measures analysis was run to analyse TCT, THT and the mental demand component of workload. The within-subject factors were
auditory, haptic and visual feedback, all at two levels, while condition order was treated as a between-subject factor.

4.1. Trial completion time (TCT)
The repeated measures analysis performed on TCT revealed a main effect of haptic feedback ($F_{(1,24)} = 5.591, p < 0.05$) (figure 2). The mean TCT for the conditions without haptic feedback (i.e. none, auditory, visual, and auditory/visual) was 60.58 s, while the mean for the conditions with haptic feedback (i.e. haptic, haptic/auditory, haptic/visual, and auditory/haptic/visual) was 62.59 s. This reveals that participants were able to complete the task at a significantly faster rate without the presence of haptic feedback than with haptic feedback.

A significant two-way interaction was also found for TCT, between haptic and auditory feedback ($F_{(1,24)} = 7.790, p < 0.05$) (figure 3). This interaction not only shows the strong haptic main effect, but also the bimodal effect of auditory and haptic feedback combined. In the bi- and trimodal conditions where haptic and auditory feedback were combined (i.e. auditory/haptic, and auditory/haptic/visual), the mean TCT was at its highest (mean 62.76 s). When the haptic feedback was removed, leaving the bimodal condition of auditory/visual and the unimodal condition of auditory feedback, the participants were able to reduce their mean TCT (mean 59.39 s). The other marginal means displayed in figure 3 are listed in table 3.

4.2. Target highlight time (THT)
Two significant main effects were found for target highlight time: auditory ($F_{(1,24)} = 15.601, p < 0.05$), and haptic ($F_{(1,24)} = 6.053, p < 0.05$). THT increased when auditory feedback was present and decreased when haptic feedback was present. There was a mean difference of 0.329 s between the mean of the conditions without auditory feedback (i.e. none, haptic, visual, and haptic/visual) and the mean of the conditions with auditory feedback (i.e. auditory, haptic/auditory, auditory/visual, and auditory/haptic/visual) (figure 4). The haptic main effect, in comparison, demonstrated that the mean of the conditions without haptic feedback (i.e. none,

![Figure 2. Haptic main effect for TCT.](image-url)
auditory, visual, and auditory/visual) was 0.767 s, while the mean of the conditions with haptic feedback (i.e. haptic, haptic/auditory, haptic/visual, and auditory/haptic/visual) was significantly shorter, 0.689 s (figure 5).

There was a three-way interaction detected between auditory, haptic and visual feedback ($F_{(1,24)} = 4.631, p < 0.05$). The strong influences of the auditory and haptic main effects are particularly evident. To illustrate this interaction best, two two-way interactions are shown in figures 6 and 7. Figure 6 shows the interaction between visual and haptic with no auditory feedback present. The unimodal condition of haptic feedback alone (mean 0.499 s) and the bimodal condition haptic/visual (mean 0.537 s) provided the two quickest THTs of the eight conditions. Figure 7 shows the interaction between visual and haptic with auditory feedback present. As shown, the presence of auditory feedback led to an overall increase in THT. The positive influence of haptic and visual feedback on performance (figure 6) was reversed when auditory feedback was added. All the marginal means shown in figures 6 and 7 are listed in table 4.

4.3. Workload: mental demand
A significant main effect of visual feedback ($F_{(1,24)} = 4.674, p < 0.05$) was found for mental demand. When visual feedback was provided, the level of perceived mental

<table>
<thead>
<tr>
<th>Auditory</th>
<th>Haptic</th>
<th>TCT mean (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>Absent</td>
<td>61.777</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>62.430</td>
</tr>
<tr>
<td>Present</td>
<td>Absent</td>
<td>59.395</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>62.764</td>
</tr>
</tbody>
</table>

Figure 3. Haptic × Auditory interaction for TCT.
demand decreased. The mean mental demand ranking for the mean of the conditions without visual feedback (i.e. none, auditory, haptic and auditory/haptic) was 31.92. The ranking for the mean of the conditions with visual feedback (i.e. visual, auditory/visual, haptic/visual and auditory/haptic/visual) was 28.95. The scale ran low (0) to high workload (100), so the perception of mental demand decreased when visual feedback was present.

A significant three-way interaction was also found between auditory, haptic and visual feedback ($F_{(1,24)} = 10.579$, $p < 0.05$). Again, the three-way interaction will be depicted in two, two-way illustrations in figures 8 and 9. Figure 8 shows the auditory by haptic interaction with visual absent, while figure 9 shows the interaction with visual feedback present. Auditory feedback and haptic feedback were strongly
affected by the presence of visual feedback. The unimodal condition of visual feedback (mean 25.85) resulted in the lowest mental demand. When haptic/visual were combined, the mental demand ranking remained low (mean 26.95). However, when visual feedback was absent and the bimodal condition of auditory/haptic was presented, mental demand was at its highest (mean 35.70). The remaining marginal means are listed in table 5.
5. Discussion

The three dependent variables, TCT, THT and mental demand, each represent a unique aspect of task performance and indicate how multimodal interfaces can best be integrated into direct manipulation tasks such as the drag-and-drop task tested in this study.

One of the most intriguing aspects of the results is the contrasting findings pertaining to TCT and THT. One contrast occurred for the haptic feedback condition. As shown in figures 2 and 5, the presence of haptic feedback was significantly detrimental to TCT yet significantly enhanced THT. A second contrast (figure 4) illustrated that auditory feedback was significantly detrimental to THT. While there was no main effect due to the auditory condition on TCT, the means indicate an improvement to TCT in the presence of the auditory condition. To interpret these contrasting results properly, a closer examination of the performance measures is necessary. TCT and THT are related measures that represent distinctly different aspects of feedback. TCT represented the time it took to complete one

Table 4. Target highlight time—marginal means.

<table>
<thead>
<tr>
<th>Auditory</th>
<th>Haptic</th>
<th>Visual</th>
<th>THT mean (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>0.656</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>Absent</td>
<td>0.565</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>Absent</td>
<td>0.499</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>Present</td>
<td>0.537</td>
</tr>
<tr>
<td>Present</td>
<td>Absent</td>
<td>Absent</td>
<td>0.900</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>Absent</td>
<td>0.948</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>Present</td>
<td>0.903</td>
</tr>
<tr>
<td></td>
<td>Present</td>
<td>Present</td>
<td>0.820</td>
</tr>
</tbody>
</table>

Figure 8. Auditory × Haptic interaction when Visual is absent for Mental Demand.
entire trial. This broad measure encompasses many elements of task performance. In this study, TCT represented overall task performance. In other words, performance times such as time to locate file icons, travel time, positioning time and time to locate folder icons. THT, on the other hand, more succinctly represented the effect of the feedback because the measurement of THT began when a file was positioned over its target and concluded when the file was released. A true measure of feedback is how quickly it conveys meaning to the user. THT measured precisely that element of response time. With this distinction between TCT and THT in mind, the two contrasting results can be discussed.

The beneficial effect of haptic feedback, shown by the specific measure of THT, was expected. Haptic feedback is known to be one of the fastest ways to transmit information (Nelson et al. 1990). The fast response time associated with haptic feedback cannot be modified because it is hard-wired (Sanders and McCormick 1993). Since THT was a strong indicator of the effect feedback had on the

![Figure 9. Auditory × Haptic interaction when Visual is present for Mental Demand.](image)

Table 5. Mental demand ranking—marginal means.

<table>
<thead>
<tr>
<th>Auditory</th>
<th>Haptic</th>
<th>Visual</th>
<th>Mental demand mean ranking</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>Absent</td>
<td>Absent</td>
<td>32.188</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Present</td>
<td>25.859</td>
</tr>
<tr>
<td>Present</td>
<td>Absent</td>
<td>Absent</td>
<td>30.859</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Present</td>
<td>26.953</td>
</tr>
<tr>
<td>Present</td>
<td>Absent</td>
<td>Absent</td>
<td>28.594</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Present</td>
<td>32.188</td>
</tr>
<tr>
<td>Present</td>
<td>Absent</td>
<td>Absent</td>
<td>35.703</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Present</td>
<td>30.547</td>
</tr>
</tbody>
</table>

Multimodal feedback
participants’ actions, the shorter the THT the more effective the feedback. However, the results did not demonstrate the same effect for TCT. This was, at least in part, due to the all-encompassing nature of the TCT performance measure. Aspects such as time to locate file and folder icons, travel time, positioning time, as well as THT were included in the TCT measurement.

While the presence of auditory feedback was beneficial for TCT, it was significantly detrimental to THT. The reversal of the effect of auditory feedback may have resulted from the detail at which TCT and THT measured performance, more specifically the length of time measurement. The mean TCT for the experimental task was around 60 s, whereas the mean THT was 0.80 s. Recall that auditory feedback was presented by an earcon consisting of musical tones. Unlike a single beep, the earcon lasted for approximately 1 s. Although not restricted by the playing of the earcon, the participants tended to wait until the earcon had finished playing before dropping the file into the folder. Consequently, the THTs were significantly higher when auditory feedback was presented. There were also several other interesting findings in the results.

The interaction between haptic and auditory feedback for TCT revealed that both unimodal (auditory) and bimodal (auditory/visual) conditions yielded the quickest times to complete the trial. This was because the auditory feedback earcon enabled participants to perform the task quickly in the auditory condition as well as in the auditory/visual feedback bimodal condition. The unimodal condition of visual feedback, although the most familiar to the participants, produced slower TCTs than the unimodal (auditory) and bimodal (auditory/visual) conditions just discussed. As gauged by this measure, this result substantiates the claim that visual feedback alone may not always ensure the fastest performance.

Further information was provided by the auditory by haptic by visual interaction for THT. Unimodal conditions of haptic and visual feedback and the bimodal condition of haptic/visual feedback produced the quickest THTs. The remaining four feedback conditions had an opposite effect on THT. The unimodal condition of auditory feedback, and the bimodal conditions of auditory/haptic and auditory/visual, along with the trimodal condition of auditory/haptic/visual, resulted in slower THTs.

Analyses involving one-sample $t$-tests were conducted in order to generate comparisons between the conditions investigated in this study on THT. The results of the analyses are summarized in table 6. From these results, figure 10 was constructed as a representation of the feedback conditions that are most effective and least effective on target highlight time.

The results for the auditory by haptic by visual interaction for mental demand showed that a unimodal condition (visual feedback) and a bimodal condition (haptic/visual) led to a low level of mental demand. After visual feedback and haptic/visual feedback, the unimodal condition of auditory feedback produced the lowest mental demand ranking. The bimodal condition of haptic/visual feedback, which resulted in decreased mental demand, also yielded one of the fastest performance times as measured by THT. This consistency between performance and perception of performance for the bimodal condition of haptic/visual is quite compelling and warrants further development.

Some bimodal feedback was shown to yield performance times and mental demand scores at least as good, or better than, some unimodal conditions. These findings are in accordance with what other multimodal researchers have found (Campbell et al. 1999). Overall, these results demonstrate the benefits of multimodal
feedback. The results reported here are unique because a study of this type investigating bi- and trimodal feedback has not been reported in the literature to date. Therefore, the general finding that bimodal conditions enhanced performance is quite persuasive.

6. Conclusions
The experimental task was a direct manipulation task that required target matching, much like tasks performed in Windows®-like environments. Whether a user has done something correctly or incorrectly, the purpose of feedback in these environments is usually either to validate an action or to assist a user in reversing an action. As a result, a great deal of emphasis was placed on the measure of THT when formulating recommendations for the use of multiple modalities. For this particular task, the bimodal condition of haptic/visual feedback was quite beneficial
to user performance. Not only did it lead to enhanced performance, but also to a positive perception of performance, as measured by mental demand.

Although the exact combination of modalities needed to produce maximum performance is task- and user-dependent, several general recommendations can be made based on these findings.

- Haptic and visual feedback are more beneficial when used alone as well as in combination with each other when compared with other uni-, bi- and trimodal conditions.
- Temporal nature of auditory feedback alone and in combination can hinder performance. Observations confirmed that participants waited for the auditory feedback to conclude before continuing to complete the task.
- Bimodal feedback, appropriate to the task, offers performance benefits while not increasing workload.

As a result, designers should include multiple types of feedback in GUIs. Multimodal feedback allows users the ability to expand upon traditional means of human–computer communication. This is particularly essential when a traditional type of feedback, such as visual, is overused. However, caution must be used when developing the specific combination of modalities because not every bimodal feedback condition affects performance in the same way. Special consideration must be paid not only to the task, but also to the requirements that the task places on the user. If the user is unable to meet these requirements, redundancy of feedback is potentially invaluable. Ultimately, the continued dissemination and exposure of new multimodal interfaces can effectively and efficiently broaden and fortify the communication channels between humans and computers.

Acknowledgements
This research was made possible through funding awarded to Julie A. Jacko by the Intel Corporation and the National Science Foundation (BES-9896304). It also constituted a portion of Holly S. Vitense’s dissertation research.

References


