Negative Feedback for Small Capacitive Touchscreen Interfaces: A Usability Study for Data Entry Tasks

Sarangi P. Parikh and Joel M. Esposito

Abstract— Touchscreen technology has become pervasive in the consumer product arena over the last decade, offering some distinct advantages such as software reconfigurable interfaces and the removal of space consuming mice and keyboards. However, there are significant drawbacks to these devices that have limited their adoption by some users. Most notably, standard touchscreens demand the user's visual attention and require them to look at the input device to avoid pressing the wrong button. This issue is particularly important for mobile, capacitive sensing, non-stylus devices, such as the iPhone where small button sizes can generate high error rates. While previous work has shown the benefits of augmenting such interfaces with audio or vibrotactile feedback, only positive feedback (confirmation of button presses) has been considered. In this paper we prototype a simple system that provides negative vibrotactile feedback. By negative, we mean the feedback that is generated when an inactive or ambiguous part of the screen, such as the area between two buttons, is touched. First, we present a usability study comparing positive and negative vibrotactile feedback for a benchmark numerical data entry task. The difference in performance was not statistically significant, implying negative feedback provides comparable benefits. Next, based on the experimenter's observations and the user's comments, we implement a multi-modal feedback strategy – combining complementary positive audio and negative vibrotactile signals. User testing on a text entry experiment showed that with multi-modal feedback, users exhibit a (statistically significant) 24% reduction in corrective key presses, as compared to positive audio feedback alone. Exit survey comments indicate that users favor multi-modal feedback.

Index Terms—Operator interfaces, Hardware/software interfaces, Commercial robots and applications, Usability testing

1 Introduction

In recent years, touch screen technology has progressed rapidly and has increased in popularity with developers and consumers alike. Touchscreen devices, in the forms of shopping kiosks, ATMs, tablet PCs, and smart phones, are now prevalent in everyday life. They offer several advantages. First, a single touch screen can replace an LCD display, a conventional keyboard, and mouse or joystick; drastically reducing the physical footprint of the user console. Second, for novice users they are more intuitive because touch interfaces can be configured to provide a one-to-one correspondence between the available input options and the user input device; eliminating the extra unused buttons seen on some keypad interfaces, or the need to use awkward button combinations (e.g. control-alt-delete). Finally these interfaces are software reconfigurable enabling the programmer to present the user with a custom designed layout suited to each phase of the task, which maximizes efficiency and usability.

Conversely, the problems associated with touch screen use are also becoming more apparent, as they are used for a greater variety of tasks. First among these is the lack of intrinsic haptic feedback. Rapid touch typing on a traditional keyboard is facilitated by the contoured button surface and mechanical resistance of the keys. In contrast, non-augmented touchscreen buttons demand constant visual attention to avoid pressing the wrong button. A second issue is that the display and input device are typically collocated, which may not be ergonomic for extended use.

The relative weighting of the pros and cons is specific to both the user and the nature of the task at hand. For example, the advantages of touchscreens seem to outweigh the disadvantages for novice users, whom will likely prefer a more intuitive, task customized touchscreen. Similarly, the touchscreen's ergonomic drawbacks will likely go unnoticed for tasks which have brief durations or require very few interactions with the touchscreen.

In this paper we focus on improving touchscreen experiences for *experienced users*, executing *demanding tasks*. We define the *experienced user* as someone who expects to perform the task frequently, for extended periods. As such they are willing to tradeoff "easy to learn" for performance, similar to those touch typists [1]. We define a *demanding task* as one that: (1) requires users to divert visual attention away from the input device for performance monitoring; and (2) requires a high number of interface interactions per unit time.

In this paper, we explore the idea of *negative* vibrotactile feedback for small, capacitive touchscreen interfaces.

S.P. Parikh, and J.M. Esposito are with the Systems Engineering Department, United States Naval Academy, 105 Maryland Ave, Annapolis, MD 21402 USA(phone: 410-293-6111; fax: 410-293-2215; e-mail: sparikh, esposito@usna.edu).

Manuscript received 1 July2011.

By negative, we mean that the interface should provide feedback when the user does something potentially erroneous or ambiguous; in contrast to positive feedback which typically confirms each interaction. We hypothesize that for the experienced user, performing demanding tasks, this type of feedback can improve performance and ergonomics; while being less irritating than a constant stream of positive feedback. Section 2 reviews related work on touchscreens and vibrotactile feedback; and explicitly states the research questions addressed in this study. Sections 3 - 4 describe a controlled usability study performed to (in)validate these hypotheses and discusses lessons learned.

Section 5 takes the lessons learned from the initial study and attempts to design an improved small, capacitive touchscreen interface that combines negative and positive feedback. Section 6 presents the results of a more realistic usability study comparing the new feedback mode with the mode presently available on a popular consumer product (Apple's iPhone). Concluding remarks and the relationship of these findings to existing knowledge are presented in Section 7.

2 Related Work and Research Questions

2.1 Touchscreens: Benefits and Drawbacks

There are both benefits and drawbacks to touchscreens. One of the most obvious benefits is the ability to customize the keyboard layout as required by the user or the task. Potter et al. conducted a series of studies that have impacted the design of touch screens in order to make them comparable to conventional input devices [2]. Sears et al. studied the relationship between button size, finger size and typing speed [3]. An additional advantage is that touchscreens also eliminate the need to have additional accessories such as keyboards and mice [4], resulting in smaller devices.

While touchscreens are invading the handheld device market, there are a few drawbacks that limit their adoption. First, and foremost the lack of haptic feedback negatively impacts an experienced user's typing speed [5, 6, 7] as compared with traditional keyboards. Without haptic feedback users often have to repeatedly look at the input device to avoid pressing the wrong button, they cannot find the home keys quickly, and engage in more frequent error checking.

In addition, it has been shown that smaller key sizes, common on mobile devices, result in higher error rates. Button sizes under 22 mm [8], [9] increase error rates – consider that the average thumb size for adult males and females are 22.9 mm and 19.1 mm respectively [10]. Resistive touch sensing technology, which measures pressure, offers some options for working with small buttons. Lee [11] found that users presented with small 5 X 8 mm buttons– the same size as the iPhone 4 in portrait mode—prefer to use fingernails or non-specialized mechanical styluses to improve typing speed and accuracy. Fingernail typing is not possible on a capacitive touchscreen, and a specialized stylus is required. Also because a range of force levels can be detected with resistive sensing, single-target selection aids can be easily implemented such as takeoff cursors [4] and mouse over events [12], which dramatically improve performance. Such methods are not well suited to high frequency interactions such as typing. In addition, they are difficult (though possible) to implement on a capacitive touchscreen, because "touch" is typically treated as a binary event (i.e. a click).

Finally, there are ergonomic drawbacks associated with the fact that a single screen typically serves as an input *and* display device. Repeatedly touching a screen at eye level can be tiring for the arm muscles – a well-established condition known as "gorilla arm" [13]. Similarly, studies have shown that continuously looking down at table mounted or hand held screens can result in neck fatigue [14]. These effects are amplified during the extended interactions heavy users engage in.

2.2 Touchscreen Technologies

There is much work on haptically augmenting the touchscreen experience. Bau et al. [15] developed TelsaTouch which uses high fidelity tactile feedback, generated by electro-vibration to simulate a variety of textures. Similarly, Levesque et al. [16] showed that when imperceptible high-frequency vibrations are used to generate variable friction on a touch surface user engagement increases. Jansen [17] introduced MudPad which uses an array of magnets and a layer of magneto-rheological fluid to induce a variety of textures. Poupyrev et al. [18] presents Lumen which uses shape memory alloys, while Harrison et al. [19] uses inflatable buttons. Hoggan et al. [20] used two higher resolution vibratory actuators to provide more nuanced feedback. A separate set of work focuses on generating haptic signals via a stylus, rather than through the screen directly. Examples of actuated pens include the Haptic Pen [21] and Anoto Functionality [22]. Other stylus-based approaches use the actuators on the display to respond to a specially instrumented stylus [23], [24].

At this time none of these technologies have penetrated the mobile consumer device market. Possible reasons include lack of technical maturity, cost, form factor or power requirements. At present, haptic feedback on consumer devices is typically generated by a simple, low resolution, vibration motor.

2.3 Performance Enhancement via Vibrotactile Feedback in Mobile Consumer Devices

Lee et al. [11] compared non-augmented touchscreen performance to various positive feedback types: audio, vibrotactile and combined audio/vibrotactile. We use many of the same typing performance metrics in our study (e.g. keystrokes per character, characters per second, etc.). They established that the addition of feedback significantly improved performance as compared to a standard touch screen. However, there was no significant difference between the feedback modalities (audio vs. haptic). We exploit this observation later when we introduce multi-modal feedback.

Hoggan et al. [18] used a resistive touch sensing mobile phone to generate a variety of Tactons – distinct vibro-tactile signals in response to events such as touching the home keys, mouse-over button and button clicks. The feedback is generally *positive* in that it indicates when an event has occurred. They established that while a traditional (hard) keyboard is superior in terms of performance, the tactile feedback significantly improved text entry as compared to a non-augmented soft keyboard. Further, the Tactons reduce the users' perceived workload. They also found these improvements to be statistically significant. They employed a similar workload assessment tool to the one used in this paper – a NASA TLX Survey augmented with a distraction dimension.

Leung et al. [25] implemented a similar variety of haptic signals, except that they were customized for different GUI elements such as progress or scroll bars. Performance improvements varied greatly with the element type. The participant exit interviews point out two issues. First, users have a preference for "moderate and selective use" of haptic feedback. Second, haptic enhancements were generally regarded as more useful when they were used in a way to prevent a the user from having to devote visual attention to a task.

Only a couple of works consider negative feedback. Hogan et al. [18] implemented a finger-tip slip Tacton, triggered when one drags their finger off of a soft button. They did not individually assess the effects of each Tacton so its possible benefits were not separately shown. Similarly Haptic Button Edges [26] were rated favorably by users, and are implemented in the Nintendo Wii. Neither of these approaches considers small capacitive touchscreen devices.

2.4 Research Questions

Research has shown that the addition of both audio and vibrotactile feedback can improve typing performance. However all of these studies only implement positive feedback, as discussed in Section 2.3.

> ***Q1:*** *Could augmenting touchscreen interfaces with negative vibrotactile feedback result in levels of performance comparable to those observed using positive vibrotactile feedback?*

While previous work suggests that many users prefer positive feedback to non-feedback augmented touchscreens, some users complain that the constant stream of feedback is annoying or distracting.

> ***Q2:*** *Could the addition of negative vibrotactile feedback to a touchscreen interface improve a user's experience as compared with positive vibrotactile feedback?*

3 Design principles and Prototype

Modifying the iPhone: As discussed in Section 2.1, the performance drawbacks of touchscreen interfaces are most exaggerated when buttons are small. In addition, many of the performance mitigation strategies are only effective for resistive touch sensing (finger nail typing, stylus use, etc).

Considering the popularity of capacitive mobile touchscreen devices designed to work with finger-based typing (e.g. iPhone, Droid, Motorola Milestone, Google Nexus one), and easy access to a large population of users, we decided to investigate enhancement strategies specifically targeted at this type of consumer device.

We chose to modify the iPhone 4 according to the following set of design principles in our prototype.

The proposed design should not require any special purpose hardware. We wanted to investigate how *existing* mobile device, such as the iPhone, could be enhanced via software; rather than suggest manufacturers add special purpose hardware, or that consumers make after market modifications to their devices.

An input device should not require constant visual attention. In [25] users noted that the most useful haptic enhancements free them from having to look at the input device. Casual observation suggests, remote control plane pilots or professional typists rarely look at their input devices.

Frequent feedback may be irritating to the user. In [25] users expressed a preference "for selective or moderate haptic feedback". Frequent sensations can be stress inducing [27]. Anecdotal evidence suggests certain groups of users choose to turn off the text entry audio feedback feature on their iPhone because they find it annoying.

3.1 Interfaces

Specifically, an iPhone 4 (capacitive, 58 X 115.2mm touch screen) was used throughout this study. As seen in Fig. 1, a numerical keypad-style interface was developed for the first set of experiments, involving numerical data entry. We used a 3rd party GUI development environment to give us tighter control over the layout and data logging. However, this interface closely matches a standard numerical keypad in terms of button arrangement. The key dimensions (4 X 6 mm) and spacing (2 mm vertical, and 1 mm horizontal) match that of the native iPhone QWERTY keyboard (see Fig. 2). The interface is locked in "portrait mode" similar to the native keypad dialing interface. Each key press generates a character on a separate display screen. There is a "delete" key available which allows for corrections. Pressing "enter" records the result and advances to the next entry. Note that the number keys were "padded" with additional keys around the border of the interface – creating an interface in which all the target keys have an equal number of neighboring keys (8). Without this padding, error rates are more likely to vary based on the number being entered. For example typists are more likely to generate an incorrect character when entering the number sequence "555" vs. "999" because "5" (a center key) has 8 neighboring keys while "9" (a corner key) only has 3.

Fig. 1. A screenshot of the graphical interface used for the numeric data entry task.

It is important to note that on the iPhone the "button" includes both the rectangular "visible button" (shaded area in Fig.2), as well as one half of the spacing surrounding the button (white area enclosed by dotted outline in Fig. 2). It is important to note that the dotted lines in the figure are for illustrative purposes only – the outer button boundary is not visible on the interface. Therefore, when the user touches this area between two buttons, an action corresponding to the nearest button label is generated.

Using this visual interface, we implemented two different vibrotactile feedback modes. *Positive vibrotactile feedback*, means a vibration is generated each time a button is pressed (either white or shaded area enclosed by dotted outline in Fig. 2). *Negative vibrotactile feedback* means a vibration is generated when the individual presses the area between the "visible buttons" (white area only in Fig. 2). Note that positive and negative feedback only differ in terms of the vibration feedback they provide. There is other no difference in terms of button size, visible button size, or action generated. Each time feedback is required, we generate a 250Hz vibratory signal for 150 milliseconds.

To give us more control over the signals, the vibrations are generated by an external circuit rather than the internal iPhone actuator. A LG Innotek Mini-pancake Vibrating Motor (10 mm diameter by 3 mm thick) adhered to the center of the backside of the iPhone, which is very similar in both "feel" and capability to the native iPhone motor. From a user's perspective the only perceptible difference from a standard iPhone, was a pair of ultra-thin wires extending from the top of the casing. The wires are long enough not to impose any constraints on how the device is held and flexible enough so as not to generate any noticeable mechanical resistance when the device is moved.

4 Experiment I: Numeric Data Entry, Positive vs. Negative Vibrotactile Feedback

4.1 Task Description

Users are seated at a desk and a sheet of paper with 15 entries printed in large font is placed on a paper stand next to them (the *desired list*). Each entry is 10 characters, representing a telephone number. Users must type in the characters using the touchscreen interface and press enter after each entry. The results are displayed on a full size monitor in front of them. While a text copy task is less realistic than a text creation task, it does provide a quantitative way to measure accuracy.

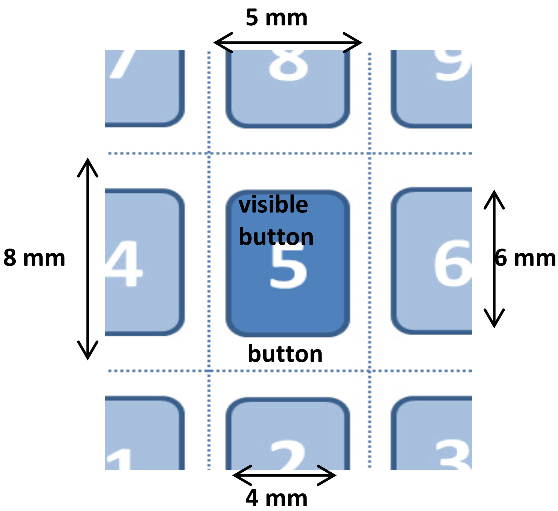


Fig. 2. Button dimensions.

The paper stand, and the monitor are placed in fixed positions on the desk across all participants and trials. The workstation layout is meant to closely mimic the ergonomic set-up a data entry technician might use when touch typing on a hard keyboard.

4.2 Methodology and Procedure

A pre-experimental questionnaire was used to screen for experienced users. We looked for individuals that use a touchscreen interface more than 5 times a day. A total of 12 experienced users volunteered for this study (all right-handed, 5 female, between 20-37 years old). The individuals in our study averaged 36.6 uses per day, typically on a Smartphone or iPad. Human subject testing approval was obtained from the local institutional review board (HRRP Approval #: USNA.2011.0004-IR-EP7-A).

To begin, the participant were seated at the desk and instructed to adjust monitor height, seating posture, and the manner in which the device is held, until they were comfortable. Although all of our participants were proficient in touchscreen usage, individuals were asked to train on the interface with each feedback mode, as long as they liked, until they felt comfortable (mean training time 2.2 minutes). This insured that the user is exposed to each feedback modality before beginning any of the experiments and helped mitigate learning effects.

Instructions were read aloud to the users before they began the experiments. They were told to enter the desired list of telephone numbers as quickly and as accurately as possible. Users were informed that they were being timed; and that the timer began at the first keystroke, and that "enter" stoped the timer and advanced them to the next number.

Post experimental surveys were also conducted after the individuals used each numerical keypad mode. At the end of all the experiments, an exit survey was completed which allowed the user to rank the two feedback modes in several categories and provide comments.

In this fully counter-balanced, within-subject experimental design, every individual completed the task twice – once with positive vibrotactile feedback; and once negative vibrotactile feedback. Half of the subjects started with the positive feedback mode while the other half began with the negative feedback mode. In total, each participant's experience lasted approximately 20 minutes.

4.3 Data Entry Performance Metrics

For each text entry task, we investigated a few, of many possible, quantitative and qualitative measures of an individual's performance and experience, including:

1. Accuracy;
2. Keystrokes per character (KSPC);
3. Keystrokes per second (KSPS);
4. Subjective Workload; and
5. Exit Survey.

Accuracy: An entry is considered correct if the text on the display screen exactly matches the desired text at the time when the user presses the "enter" key. The accuracy is the number of correct entries divided by the total number of entries (15) – ideally 1. Note that, in this performance measure, users are not penalized for incorrect keystrokes, provided corresponding corrections are made before the enter key is pressed.

Keystrokes per character (KSPC): This measures the total number of keystrokes and divides it by the number of characters required to generate an accurate entry (10, not including the "enter" key). The ideal value is 1. For example, a 10 digit number that was entered with one error and one correction would score 1.2 KSPC. Note that this measure is only meaningful for correct entries.

Keystrokes per second (KSPS): The number of keys pressed per second, is one measure of typing speed. Unlike characters per second, KSPS does not penalize correction strokes. We chose to use KSPS since the number of correction strokes is already encoded in KSPC. Note, we only compute this quantity for correct entries.

Subjective Workload: A modified version of the NASA TLX survey [28], was given to participants upon completion of their tasks. They rated their experiences along the following dimensions. The italicized illustrative descriptions of each dimensions were provided to the participants to aid them in their evaluation.

Mental Demand: (1=Demanding,…10=Easy) *E.g. need to keep data in memory; interface confusing.*

Physical Demand: (1=Demanding,…10=Easy) *E.g. required physical movement (eyes, hands, head), coordinated finger movements or applying a force.*

Frustrating: (1=Stressful, …10=Relaxing) *E.g. consistently pressing the wrong key that is nearby; didn't realize when a mistake was made; interface slows me down.*

Distracting: (1=Annoying,…10=Unnoticeable) *E.g. anything that annoys or irritates while trying to perform the task; something you may try to ignore or block out.*

Task Suitability: (1 = unusable…10 = ideal)

Overall choice: (1 = most preferred…10 = least preferred)

Exit Survey: Finally, participants were asked to provide an ordinal ranking of each feedback type; to compare the feedback types to standard touchscreens and were provided with an opportunity to comment on the experience.

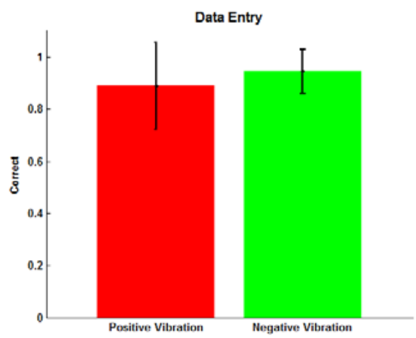
4.4 Data Entry Performance Analysis

The following table reports the mean (and standard deviation) for each of the performance measures and feedback types; while Figs. 3-5 depict the mean and standard deviation in graphical form.

TABLE1
MEAN (STANDARD DEVIATION) OF PERFORMANCE FOR NUMERICAL DATA ENTRY TASK: POSITIVE *VS.* NEGATIVE VIBROTACTILE FEEDBACK

Feedback Type	Accuracy	KSPC	KSPS
Positive	0.89 (0.17)	1.24 (0.25)	0.73 (0.07)
Negative	0.94 (0.08)	1.17 (0.11)	0.74 (0.08)

Fig. 3. Correct entries for numerical data entry task.



The means for negative vibrotactile feedback indicate a modest performance benefit in terms of Accuracy and KSPC. However, a t-test showed no significant main effect for feedback type: Accuracy ($F_{1,11}$=1.06, p =0.314); KSPC ($F_{1,11}$ =0.79, p =0.383); KSPS ($F_{1,11}$ =0.26, p = 0.615).

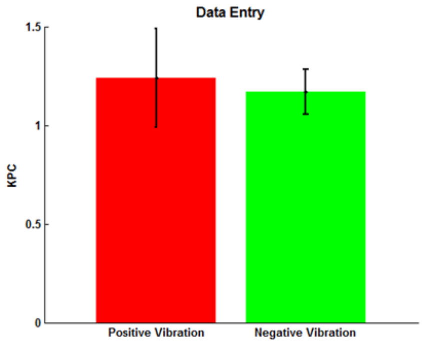


Fig. 4. Keystrokes per character for numerical data entry.

Contrary to the performance measures, the results of the NASA TLX survey (Fig. 6) suggest participants prefer the positive vibrotactile feedback. However, a t-test did not find any significant effects ($p > 0.05$) for feedback type.

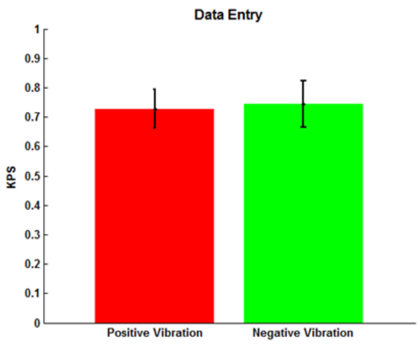


Fig. 5. Keystrokes per second for numerical data entry.

According to their exit surveys, 10 of 12 participants preferred the positive vibration feedback over negative feedback. Most indicated in the comments section that they felt the consistent stream of feedback "helped them maintain their pace". A few noted that the negative feedback "helped them catch their mistakes quickly". Note on average, the participants only activated the negative vibrotactile feedback once every 20 key presses. Interestingly, 10 of 12 thought the negative feedback was still an improvement over non-feedback augmented touchscreen interfaces they have had prior exposure to. Overall, users complained about their lack of familiarity with the numerical keypad interface, frequently pining for a QWERTY style button layout (number keys across top row) similar to that found on their smartphones – claiming that our interface layout caused them to type slower than normal. Finally, 11 of 12 users specifically inquired if it was possible to hold the interface in landscape mode. All users naturally employed one handed typing with their dominant hand.

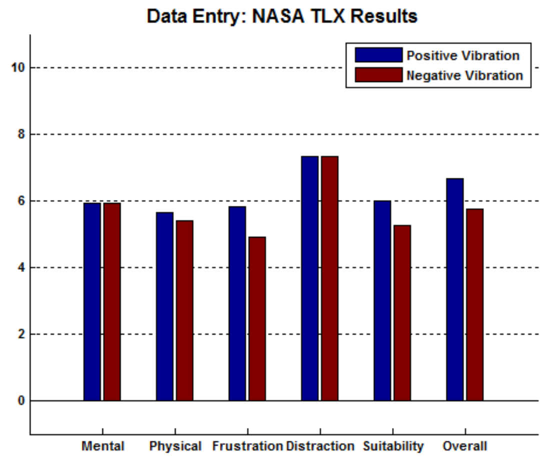


Fig. 6. NASA TLX results for numerical data entry task.

4.5 Discussion of Negative vs. Positive Feedback

Overall, we found no statistically significant differences between positive and negative vibrotactile feedback, for the interfaces we considered. Reflecting on the questions posed in Section 2, our conclusions are as follows.

1. Yes, the performance levels recorded when using negative vibrotactile feedback are statistically indistinguishable from those with positive feedback..
2. No, there is no statistically significant improvement in the user experience as indicated by the TLX survey. In fact, the TLX results for positive and negative vibrotactile feedback are statistically indistinguishable.

While there were no statistically significant differences, the observations of the participants and the experimenters raise some interesting insights. First, there is a dichotomy between performance and perception. While performance levels are indistinguishable; on their exit surveys, users suggest that positive feedback helps them maintain "speed" or "pace". Most of the comments regarding negative feedback relate to "being informed of mistakes quickly".

When observing the participants, one can see that their visual attention is almost exclusively focused on the input device as they type. Visual attention to the display screen was infrequent. Most users looked at the desired text once, then began typing. Most looked up to check their entered text on the display screen once after the number was fully entered, before pressing the "enter" key. Several others checked their text every 3 or 4 characters. Despite all having ample experience with touchscreens and data entry, none continuously monitored the desired text or their performance, the way a touch typist would.

Finally, we noted that the typing speeds reported in Section 4.4 (mean of 0.73 KSPS) was nearly half the speed reported in the literature on haptically augmented touch screens [11],[18], even once the difference in speed metrics are accounted for. Based on the exit survey comments, users believe their typing speed would dramatically increase if they:

- were permitted to hold the device in landscape orientation;
- typed words instead of numbers; and
- were given a QWERTY style button layout.

5 Revised Interface and Feedback Design

Using the observations from our first experiment, we redesigned our interface, both in terms of layout and feedback. In the new layout the button dimensions are identical to those used in the previous experiments but they are laid out in a QWERTY keyboard arrangement (see Fig. 7). The footprint and spacing replicate the native iPhone 4 keyboard. Additionally, we now allow the individuals to hold the interface in "landscape mode." The goal of these two changes was to promote higher typing speeds, where the benefits of negative feedback might become more apparent.

More importantly, the feedback modality was changed as well. We implemented a multi-modal feedback design that attempts to combine the unique benefits of positive and negative feedback, as reported on the exit surveys. As discussed in Section 2 of [11], there is no statistically significant difference between positive audio feedback and positive vibrotactile feedback. Thus, in our revised interface, users are provided audio feedback (a 50 millisecond "click") to confirm each button press, while vibrotactile signals (100 milliseconds at 250Hz) indicate that an ambiguous or potentially erroneous region of the interface was activated – the area between two buttons. While positive audio feedback – the standard feedback type on Apple's iPhone – may help users feel they type faster; negative feedback could potentially augment their experience by quickly alerting them to possible errors. We hypothesized that this approach could lead to reduction in the number of error correction strokes required.

6 Experiment II: Text Entry with Multi-Modal Feedback

6.1 Task Description

We modified the data entry task so that the list of desired entries contains commonly used 10 letter words, instead of randomly generated numbers. There were two motivations for this change. First, the exit survey suggested that participants may type faster – possibly leading to more frequent key misses. Second, we speculated that they may check their desired text less frequently, since words are much easier to remember than numbers, which are memorized and then sometimes checked in 3-4 digit groupings. Finally, we note that typing words may be a more representative task for mobile touchscreen devices.

6.2 Methodology and Procedure

The pre-experimental questionnaire that was used to screen users is identical to that described in Section 4.2. The same instructions, procedures, practice opportunities, subjective workload survey and exit survey described in Section 4.2 were used.

Fig. 7. The QWERTY interface.

A different set of 12 qualified individuals (1 female, 1 left handed male, between 21-38 years old), volunteered for this study. The participants averaged 35.2 touch screen data entry operations per day. Human subject testing approval was obtained from the local institutional review board (HRRP Approval #: USNA.2011.0004-IR-EP7-A).

This study is a fully counterbalanced, within subjects design, comparing the multi-modal feedback in Section 5 to the standard positive-audio feedback used on the native iPhone interface. Half of the subjects started with the multi-modal feedback; the other with the positive audio condition.

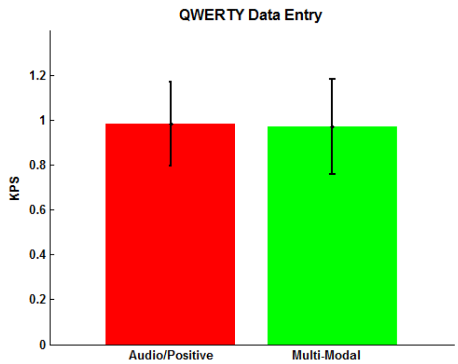


Fig. 8. Keystrokes per second for the text entry task.

6.3 Performance Analysis

We measured the performance using the same typing metrics introduced in Section 4.3. The table below reports the means (and standard deviations). See Figs. 8-10.

TABLE 2
MEAN (STANDARD DEVIATION) PERFORMANCE FOR THE TEXT

Feedback Type	Correct	KSPC	KSPS
Positive Audio	0.95 (0.10)	1.33 (0.13)	0.98 (0.19)
Multi-Modal	0.94 (0.07)	1.25 (0.11)	0.97 (0.21)

ENTRY TASK: POSITIVE AUDIO VS. MULTI-MODAL FEEDBACK

Here we found a significant effect for interface type on Keystrokes per Character ($F_{1,11}$ =5.42, p =0.025). The remaining performance metrics did not exhibit statistically significant differences: Correct ($F_{1,11}$=0.12, p =0.730). Keystrokes per Second ($F_{1,11}$ =0.07, p = 0.80).

Fig. 9. Fraction of Correct entries for the text entry task.

The subjective workload analysis (Fig. 11) suggests that users preferred the multi-modal feedback along all dimensions except for "distraction"; but a t-test showed that none of these quantitative differences were statistically significant.

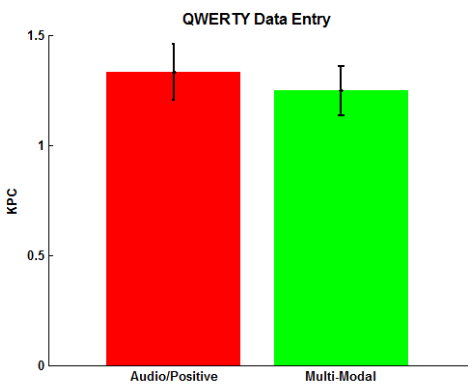


Fig. 10. Keystrokes per character for the text entry task.

Finally, according to the exit survey 10 of 12 participants preferred the multi-modal feedback over positive audio feedback; and 12 of 12 thought the multi-modal feedback was an improvement over touchscreen interfaces with no feedback. Comments on the utility of both positive and negative feedback echoed those reported in Section 4.4

6.4 Discussion of Multi-modal vs Positive Audio Feedback

By design, the combination of word entry, landscape orientation and/or the QWERTY layout results in higher KSPS across both feedback types, as compared with Experiment I. Perhaps as a consequence, KSPC was higher for this task, across both feedback types, as participants used more correction strokes; despite this, they ultimately produced more accurate results. As speculated earlier the combination of task and interface changes created conditions in which quick error correction is desirable. In addition, their mean missed key presses increased to 1.11 for both feedback types, suggesting negative feedback was activated twice as often as compared with our previous experiment (Sect. 4).

While the means suggest a 6% reduction in KSPC (or a 24% reduction in corrective keystrokes), it is interesting to point out that this reduction was not uniformly distributed over all the entries. Much of the drop results from the reduced frequency of extreme values of KSPC (>1.5). To illustrate, occasionally, when positive feedback alone was used, the first character was entered incorrectly then, without looking at the display, the participant entered the next 9 characters correctly. They realized their mistake just before pressing enter, and as a result, had to delete and then reenter the characters, which lead to 30 or more keystrokes for a single 10 character entry. These situations are prevented by the addition of negative vibrotactile feedback.

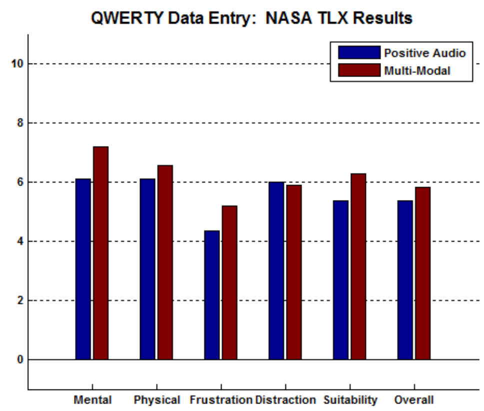


Fig. 11. NASA TLX results for the modified data entry task.

7 CONCLUSIONS AND RELATIONSHIP TO EXISTING KNOWLEDGE

In this study we investigated the effect of *negative* vibrotactile user feedback on typing tasks – an area not previously addressed in the literature. Given that few, if any, high resolution haptic actuators (e.g. [15,17,18]) have been embraced by the manufacturers of mobile touchscreen consumer devices, we choose to focus on simple software enhancement that could be implemented on existing hardware. We chose the iPhone 4 as our test platform, both due to its popularity and because it possesses design elements that were previously shown [11] to cause poor typing performance, namely: capacitive sensing, finger-based typing and small soft buttons.

For our numerical data entry benchmark task, when comparing negative feedback to previous studied positive vibrotactile feedback, we provided the answers to two research questions posed in Sect. 2.

1. Negative vibrotactile feedback resulted in typing performance that was comparable to that obtained using positive feedback.
2. Negative feedback did not result in any significant improvements in the user experience as indicated on the TLX survey.

It has been previously established [18] that when compared, users type best on a physical (hard) keyboard, followed by a touch screen with (positive) haptic feedback. The worst performance is on a standard touch screen with no haptic feedback. Our finding suggests that *type* of feedback (positive or negative) may not be a significant factor that contributes to this performance ranking.

Part of our motivation for investigating negative feedback were comments from exit surveys in another study [25] which report that users have a preference for "moderate and selective use of haptic feedback" rather than a constant stream of feedback signals. However, for our task and interface the subjective workload scores for positive and negative vibrotactile feedback were indistinguishable. A second motivation was that the authors of [23] hypothesize that the most useful haptic feedback elements prevent a user from having to look at the screen. Yet, based on our observations, negative feedback did not appear to enable touch-typing.

In general, the typing speeds we observed were considerably lower (0.74 KSPS) than those reported in the literature when using the same button sizes (1.7 KSPS in [11]) –likely due to the nature of the task. We speculate that the potential benefits of negative feedback might not be apparent at such low typing speeds. Therefore, in our second experiment we altered the task and visual interface to promote higher typing speeds (mean KSPS of 0.98) and less frequent error checking.

More importantly, we introduced a novel multi-modal feedback strategy that employed both positive audio and negative vibrotactile signals. The design was inspired by participant exit survey comments that suggest they perceived positive feedback to contribute to their "speed" while negative feedback to contribute to their "error awareness". This approach was further grounded in previous findings. [29] determined that multi-modal feedback can reduce a user's workload, while [25] reported that users occasionally had a difficult time distinguishing a large set of Tactons. Second, [11] reported that the benefits of positive feedback are identical regardless of feedback modality (audio or haptic).

Individuals using our novel multi-modal feedback experienced significant (p=0.025) 6% reduction in Keystrokes per Character (KSPC) – or a 24% reduction in corrective key strokes—as compared with those using traditional positive audio feedback alone. No other significant differences in accuracy, KSPS, or workload were detected. As a point of comparison, [11] reported that adding positive feedback (audio or tactile) to a standard touchscreen device results in a 7% reduction in KSPC. Interestingly, [18] reports that the addition of positive tactile feedback actually resulted in a 22% *increase* in KSPC, though there is some ambiguity as to how they measured KSPC in the case of incorrect entries. Based on our observations, most of the reduction in KSPC results from eliminating the negative impact of outlier entries. In such situations, a participant commits an early typing error but only detects it after the entire entry is nearly complete, resulting in a large number of corrective keystrokes. While these events are infrequent (on average once per 10 entries), they can result in 2.5 or more KSPC.

In an effort to generalize these finding, we hypothesize that negative feedback will have the greatest benefit when implemented for tasks or interfaces where:

1. users visually monitor their performance infrequently; and
2. the cost of correcting the errors increases with time.

In both of these situations, negative feedback quickly cues the user to check their performance and take immediate action if needed.

Overall, when asked to rate their experience with multi-modal feedback, 83% of the participants preferred it over the audio-positive feedback feature currently implemented on the iPhone 4.

Acknowledgement

The authors would like to acknowledge ENS Michael Martin, Mr. Joseph Bradshaw and Mr. Norman Tyson who helped develop the initial vibrotactile feedback device.

References

[1] Cooper, E. (ed.), *Cognitive Aspects of Skilled Typewriting*, Springer-Verlag, New York, 1983.

[2] Potter, R.L., Weldon, L.J., and Schneiderman, B. Improving the accuracy of touch screens: An experimental evaluation of three strategies. *Proc. ACM CHI Conference on Human Factors in Computing Systems* (1988), 27-32.

[3] Sears, A., Revis, D., Swatski, J., Crittenden, R. and Schneiderman, B. Investigating touchscreen typing: the effect of keyboard size on typing speed. *Behaviour & Information Technology, 2,* 1 (1993), 17-22.

[4] Sears, A. and Schneiderman, B. High Precision Touchscreens: Design Strategies and Comparison with a Mouse. *International Journal of Man-Machine Studies, 43,* 4 (1991), 593-613.

[5] Barrett, J. and Krueger, H. "Performance effects of reduced proprioceptive feedback on touch typists and casual users in a typing task." *Behavior & Information technology, 13*, (1994), 373-381.

[6] Lewis, J.R., Potosnak, K.M. and Magyer, R.L. "Keys and Keyboards." *Handbook of human-computer interaction*, Helander, M.G., Landauer, T.K. and Prabhu, P.V. (ed), Elsevier Science, Amersterdam (1997), 1285-1315.

[7] Findlater, L., Wobbrock, J.O., and Wigdor, D., "Typing on flat

glass: examining ten-finger expert typing patterns on touch surfaces", *Proc. ACM CHI Conference on Human Factors in Computing Systems* (2011), 2453-2462.
[8] Greenstein, J.S, "Pointing devices", in M. Helander and T.L.P. Prabhu (eds.), *Handbook of Human Computer Interaction*, Elsevier Science, Amsterdam, 1997 1285-1315.
[9] Lewis, J.R., Potosnak, K.M. amd Magyar, R.L., "Keys and Keyboards", in M. Helander and T.L.P. Prabhu (eds.), Handbook of Human Computer Interaction, Elsevier Science, Amsterdam, 1997 1285-1315
[10] Henry Drefus Associates, *The Measure of Man and Woman*, Whitney Library of Design, New York, NY 1993.
[11] Lee, S. and Zhai, S., "The performance of touchscreen soft buttons", in proceedings of ACM CHI 2009, p 309-318
[12] Albinsson, P.A, and Zhai, S., "High precision touch screen interaction", Proceeding of CHI 2003, ACM Conference on Human Factors in Computing Systems, CHO Letters 5(1) (2003), 105-112.
[13]http://www.wired.com/gadgetlab/2010/10/gorilla-arm-multitouch/
[14] Poor posture subjects a worker's body to muscle imbalance, nerve compression, Langford ML, Occupational Health Safety, 1994 Sep;63(9):38-40, 42
[15] Bau, O. Poupyrev, A.I. and Harrison, C., "Tesla-touch: elector vibration for touch surfaces", in Proceedings of UIST, 2010, p 283-292.
[16] Levesque, V., Oram, L., MacLean, K., Cockburn, A. Marchuk, N.D., Colgate, J.E., and Peshkin, M.A., "Enhancing physicality in touch interaction with programmable friction", *Proc. ACM CHI Conference on Human Factors in Computing Systems* (2011), 2481-2490.
[17] Jansen, Y., Karrer, T., Borchers, J., "MudPad: Tactile feedback and haptic texture overlay for touch surfaces", Proceedings of ITS, 2010, p.11-14.
[18] Poupyres, I., Nashida, T., Maruyama, J., Rekimoto, J. and Yamaji, Y. "Lumen: Interactive visual and shape display for clam computing", in SIGGRAPH ET 2004, p 17-25.
[19] Harrison, C. and Hudson, S.E., "Providing dynamically changing physical buttons on a visual display", In Proceedings of CHI, 2009 p 299-308
[20] Hoggan, E., Brewster, S.A., and J. Johnston, "Investigating the effectiveness of tactile feedback for mobile touchscreens", Proceedings of ACM CHI, p. 1-10 2008
[21] Less, J.C., Dietz, P.H., Leigh, D., Yerazunis, W. and Hudson, S. E., "Haptic Pen: A tactile feedback stylus for touch screens", UIST 6:(2), 291-294, 2004
[22] Anoto, "Development Guide for Service Enabled by Anoto Functionality", 2002: Anoto
[23] Fukumoto, M., and Toshaki, S., "Active Click: Tactile feedback for touch panels, in Computer Human Interaction, 2001, Extended Abstracts ACM Press.
[24] Poupyrev, I. and Okabe, M. and Maruyama, S. "Haptic feedback for pen computing: directions and strategies", in Computer Human Interaction 2004 Extended Abstracts ACM Press.
[25] R. Leung, K. MacLean, M.B. Bertelsen, M. Saubhasik, "Evaluation of haptically augmented touchscreen GUI elements under cognitive load", ICMI 2007, p. 374-381
[26] Pakkanen, T., Raisamo, R., Raisamo, J., Salminen, K.; Surakka, V.; , "Comparison of three designs for haptic button edges on touchscreens," Haptics Symposium, 2010 IEEE , vol., no., pp.219-225, 25-26 March 2010
[27] James, G. and Brown, D., "The biological stress response and lifestyle: catecholamines and blood pressure," Annual Review of Anthropology, Annual Reviews, vol. 26, pp. 313-335.
[28] Hart, S.G. and Staveland, L.E. "Development of NASA TLX (Task Load Index): Results of empirical and theorectical research", *Human Mental Workload* (1988) 139-183.
[27] Ovaitt, S., Coulston, R., and Lunsford, R., "When do we interact multi-modally? Cognitive load and multi-modal communication patterns", In Proceeding of ICMI, ACM Press 2004, p 217-220.

Sarangi P. Parikh earned a BS degree in Physics from Fairfield University, Fairfield, CT and a MS and PhD in Mechanical Engineering, in 2001 and 2005 respectively, from the University of Pennsylvania, Philadelphia, PA. She is an Assistant Professor in the Systems Engineering Department at the United States Naval Academy, Annapolis, MD. Her research interests include human-robot interaction, human-computer interactions, and interfacing with intelligent technologies.

Joel M. Esposito earned the BS degree in mechanical engineering from Rutgers University, New Brunswick, NJ, USA and the MS and PhD, in 1999 and 2002 respectively) from the University of Pennsylvania, Philadelphia, PA, USA. He is an Associate Professor in the Systems Engineering Department at the United States Naval Academy, Annapolis, MD, USA; and currently serves as the Editor in Chief of the American Society of Engineering Education's Computers in Education Journal. His research interests include multi-robot coordination, motion planning, computational tools for motion control and robotics pedagogy. Prof. Esposito is the recipient of the Naval Academy's 2010 R. A. Raouff Teaching Award.