THE IMPACT OF VIBRATIONAL DIRECTIONAL FEEDBACK ON TEXT ENTRY ON MOBILE DEVICES

DINO MIJATOVIC

This dissertation was submitted in part fulfilment of requirements for the degree of MSc Advanced Computer Science

DEPT. OF COMPUTER AND INFORMATION SCIENCES

UNIVERSITY OF STRATHCLYDE

SEPTEMBER 2016

Declaration

This dissertation is submitted in part fulfilment for the degree of MSc Advanced Computer Science of the University of Strathclyde.

I declare that this dissertation embodies the results of my own work and that it has been composed by myself. Following normal academic conventions, I have made due acknowledgement to the work of others.

I declare that I have sought, and received, ethics approval via the Departmental Ethics Committee as appropriate to my research.

I give permission to the University of Strathclyde, Department of Computer and Information Sciences, to provide copies of the dissertation, at cost, to those who may in the future request a copy of the dissertation for private study or research.

I give permission to the University of Strathclyde, Department of Computer and Information Sciences, to place a copy of the dissertation in a publicly available archive.

Yes [✓]   No [   ]

I declare that the word count for this dissertation (excluding title page, declaration, abstract, acknowledgements, table of contents, list of illustrations, references and appendices is 17598.

I confirm that I wish this to be assessed as a Type 1   2 $\boxed{3}$ 4   5 Dissertation

Signature Dino Mijatovic

Date: 29/08/2016

Abstract

One of the key disadvantages keyboards on mobile touchscreens have in comparison to standard personal computer keyboards is their lack of physical feedback. A user on a physical keyboard is made aware of immediately of when they have hit the edge of a key, making them more aware of a potential mistake. Furthermore, this feedback helps users to adjust their typing so they more often hit the centre of a key, improving their accuracy over time. These effects are not present with keyboards on flat touchscreens.

This project investigates a vibrational feedback model that seeks to mimic the feedback available to physical keyboard users. Two actuators plugged into an Android device via audio jack and attached to either side of an Android device, provide vibrational feedback that notifies users when they have hit the left or right edge of a key. This feedback takes the form of vibrations in the individual actuators, on the left and right of the device. Users holding this device will know from the actuators whether they have hit the left or right side of a key, as either the left or right actuator will vibrate. If they have hit the centre of the key, both will vibrate. An application was developed in Android where a custom keyboard was built that could take advantage of this feedback model for experimenting.

Experimentation was carried out with participants, who tested this feedback model, as well as a simpler design to mimic standard typing interfaces on Android phones, that fired both actuators whenever a key was pressed, regardless of where the key was hit. An analysis of objective data relating to participants' speed and accuracy while typing found that there was no substantial improvement in their performance when using the directional vibrational feedback. This was found to be in contrast to participants' opinions on the feedback, with a majority reporting that they found the feedback model useful, and believed their performance had increased by being able to spot their mistakes more consistently, as well as a smaller perceived amount of edge presses on the keyboard. A system that users find easy to adjust to and learn has merit, and as such further research and experimentation with this feedback model would be extremely worthwhile for improving users' speed and accuracy when typing on smartphone devices.

Acknowledgements

I’d like to thank my supervisor Mark Dunlop for his guidance throughout researching and writing this dissertation.

Thank you also to the participants who took part in the experimentation for this project.

A big thank you to Kornel Kotan and Edward McNealy, your friendship and comradery helped drive this dissertation forward.

And finally big thanks to Vedran Mijatovic and Rebecca Leith for your constant support.

Contents

Table of Images

Table of Tables

1. Introduction

Utilising a physical keyboard when typing ensures that a user's fingers get automatic feedback relating to the position of their finger on keys and their accuracy, i.e. feeling the surfaces and edges of keys. Touch screen keyboards, due to being on flat, even plane, offer no such feedback. This lack of feedback leads to a diminishing of user accuracy and speed when typing (Hoggan, Brewster, Johnston 2008). These reductions in accuracy are usually attempted to be rectified by underlying prediction engines and language models used to interpret taps – however the worse the taps the more likely auto correction errors are.

Prediction algorithms designed to improve accuracy and speed of typing have existed as far back as during the era of 12-key mobile devices (Gong, Tarasewich 2005). Predictive world completion techniques were designed in order to improve efficiency and speed when typing on these kinds of devices, as both were hindered by the small amount of keys on phones that could be used. However the nature of physical keys still meant that accuracy could be improved by providing physical feedback from key edges to the user's fingers.

With the advent and rise to prominence of touch screens, this form of physical feedback is no longer attainable, due to the flat nature of these screens. This results on "far slower and error-prone" text entry on smartphones with touchscreens than on devices with physical keys (Brewster, Chohan, Brown, 2007). To combat this, predictive entry algorithms/engines have been designed and implemented in order to improve users' accuracy and speed when typing. This is done through auto- correct functionality that is designed to detect errors in input, as well as word prediction, allowing users to enter a word into text before they have finished typing it. However these systems are not immune to error themselves, and in increase in human error leads to a decrease in their own accuracy (Dunlop and Taylor 2009).

This project followed on from the work of Dunlop and Taylor (2009), and Hoggan, Brewster and Johnston, investigating the applicability and benefits of implementing tactile feedback on mobile device text entry. Specifically, and similar to the work of Dunlop and Taylor, the use of vibrational feedback was explored, and how it could create an improvement in current typing speeds on mobile devices. The purpose of the feedback method designed and tested in this project was to investigate whether mimicking the physical feedback of key edges present in physical keyboards could improve users' accuracies. They would be more aware of their mistakes, and would improve their accuracy over time by adjusting their typing to hit less key edges. Through the use of purposeful, and directional vibration on a device, the feasibility of using such a method to guide users based on their taps (eg inform if they tapped too far left by a vibration on the left of the phone) and increase their accuracy when typing was assessed.

1.1 Research Questions, Aims and Objectives

The research questions, aims and objectives of this project were as follows:

Research Questions

"How can users' accuracy and speed with text entry on mobile devices be improved through the implementation of vibrational tap feedback on devices?"

"What will users' subjective responses be to the use of vibrational feedback for text entry, pertaining to its ease of use, and perceived usefulness?"

Aim:

Investigate the effects and benefits of vibrational tap feedback on text entry accuracy on mobile devices.

Objectives:

- Research and analyse different prediction engines used in mobile devices for implementation in the project
- Design and implement a vibrational feedback model that will utilise a prototype device designed specifically for testing in this project
- Analyse the usefulness of this model through user testing, taking into account speed and accuracy of user text entry with and without it, as well as user feedback about the performance and ease of use of the feedback model itself, as well as its effect on the users themselves
- Advocate future research on the subject.

1.2 Overview of Research and Methodology

Research was undertaken into previous methods of improving users' performance when performing text entry on mobile devices. These methods included prediction and correction algorithms, as well as tactile feedback in the form of vibrations on the device, and gave a cohesive background to use when creating and designing the requirements, and experimentation methodology of this project.

A physical prototype of device was developed and used for development and testing. A Motorola Moto G was the Android device used. Two Haptuator MkII actuators from Tactile Labs (2016) connected to the phone via its audio jack, were attached to either side. This actuators served to produce the specific directional vibrational feedback necessary for this project. If a user were to press the left edge of a key, the left edge would vibrate to let the user know, and the same holds for the right side. The effects of this feedback were meant to mimic the feedback gained from a real physical keyboard, allowing a user to spot mistakes earlier by being aware of the fact that they've not hit the centre of the key they were aiming for, and also to adjust their typing to be more centred on the keys.

Development took place in Android Java. An application was developed that contained a custom keyboard built in order to be able to produce specific feedback if the edge of a key is pressed. Also contained in this application was a predictive engine for text entry, developed using Android's Spell Checker Service. The feedback model was implemented into this application by playing specific sounds that would create vibrations so signify a centre, left edge, or right edge key press.

For experimentation, nine users tested the application to give quantitative and qualitative data on its effect on text entry. This application contained two sets of 50 phrases that would be displayed to participants when testing was carried out that they would type out. During this testing, metrics tracked included how close the users' input was to the target phrase, time taken for the input, total backspaces, and edge presses, which is what made up the quantitative data for this project. Users would type in 50 phrases using a standard, "mono" feedback, where users are made aware only of when they have hit a key, and the "stereo" feedback model, where users are made aware of their key presses, as well as when they have hit a key edge. During testing, users were given several forms and questionnaires to rate each feedback model, and to gain their opinion on the usefulness of the stereo feedback model, how easy it was to use, and how they felt it impacted their performance.

From the objective data gathered, it was concluded that there was no noticeable or substantial difference between the two feedback methods tested. The metrics gathered indicated that there was little or no improvement in typing speed and accuracy during text entry while using the stereo feedback model. However this was found to be in direct conflict with the subjective data gathered, as most users reported a positive response to the feedback model, and perceived their performance to improve when using the stereo feedback model. It was deemed that the small size of the number of participants created a small amount of data that was difficult to draw conclusions from, and that larger amount of participants would have resulted in data that would have led to more positive results.

2. Literature Review

Mobile devices, in particular mobile phones, are devices that have extremely prominent in a large variety aspects of people's lives. They provide convenient and straightforward methods of communication, accessing information, and media content, and with access to the internet, its services and functionalities, have become extremely common, useful tools.

Since the advent of these devices, a key facet of these devices and their functionality is their method of text entry and input, a facet constantly being improved on. Due to the nature of these devices being small enough to be considered handheld, this has led to smaller keys on the devices. The earliest widespread and successful key layout was the 12 key layout (Dagiene, Grigas and

Jevsikova, 2011). As pictured below, 12-key keypads have 10 number keys with letters assigned to them in alphabetical order and 2 extra keys for the "*" and "#" symbols.

Figure 1 12-key keypads

At the time of (Dagiene, Grigas and Jevsikova's report (2011), 2 variants existed of the 12-key keypad, one containing basic Latin letters, digits, and other Latin letters of a particular native language, and of some foreign languages. The other variant included alphabet letters of a particular native language, digit, and letters of some foreign languages. This allowed this input method to be tailored to a wide range of languages and character types.

As pointed out by Butts and Cockburn (2002), the devices of the time were “not naturally suited to text input”. The 12key input method had various disadvantages and drawbacks. Pavlovych and Stuerzlinger (2003) describe the frequently encountered “multitap” method found in this input model was a method in which a user pressed the corresponding key multiple times until the letter they wished they appeared, (for example, pressing ‘2’ once to enter ‘a’, twice for ‘b’, etc). As outlined by their report, a “notable difficulty with *Multitap* is entering consecutive letters that appear on the same key (a problem called *segmentation*)”. This tends to mean that a user’s typing speed is slowed down significantly as they have to utilise a “timeout”, waiting for a set of time before the cursor advances and lets the user enter the next character.

2.1 Touch Screen Usage

With the rise in popularity of smartphones with touch screens, mobile devices eventually transitioned into having QWERTY keyboards of their own. This started from earlier PDA devices, that used finger and stylus typing and eventually became commonplace in the current smartphones that dominate the handheld market. Similarly to early 12 key devices, the smaller size of these keys compared to that of a traditional QWERTY keyboard hindered users’ speed and accuracy when typing (MacKenzie, Isokoski, 2008**)**. As mentioned by **(**Sunghyuk Kwon, Donghun Lee, Min K. Chung), text entry with smaller QWERTY passed keyboards runs into problems such as “the small size of the virtual keys, absence of tactile feedback, and occlusion of virtual keys by fingers”. One distinct disadvantage these touchscreen-based keyboards have compared to their 12key keypad predecessors is the lack of tactile feedback from actual physical buttons. As put by Tinwala and MacKenzie (2010, “because the keys are physical, users feel the location of buttons and eventually develop motor memory of the device”. Due to the nature of smartphones’ flat screen, there is no way to get any definite physical feedback, and as such is not the case with modern QWERTY keyboards. The effect of this is that there is ‘an increased need to visually attend to the device. The effect is particularly troublesome if the user is engaged in a secondary task. Consequently, the high visual demand of touch input compromises the “mobile” in “mobile phone”’(Tinwala and MacKenzie, 2010). The most prevalent method of trying to tackle this issue and to increase typing speed and accuracy with this input model is to aid users with their typing through the use of supplementary algorithms for error correction and word prediction.

2.2 Dictionary Based Text Entry Models

Development and implementation of word prediction algorithms that worked by "using a single key-press per letter together with a large dictionary of words for disambiguation"(Dunlop, Crossan 2000) were attempts to rectify and improve upon the slower and sometimes inaccurate and slow typing with touchscreen keyboards. These methods were given the broad ranging definition of "Dictionary-Based Disambiguation" by Pavlovych and Stuerzlinger (2003). Prediction and word completion techniques were also joined by research into the design of keys and layouts, (Dunlop, Levine 2012) and what letters each key would represent.

Further attempts to improve text input designs of unique keyboard layouts (Oulasvirta et al 2013), analysis of different input methods and postures (Azenkhot, Zhai 2012) as well as specially designed games created to give users practice using smaller touchscreen QWERTY keyboard layouts (Rudchenko, Paek, Badger 2011). The most prevalent area of research, and admittedly most successful in improving the experience of text entry on mobile devices, is that of prediction and error detection algorithms. As noted by the results of Sunghyuk Kwon, Donghun Lee, Min K. Chung's work, users seem to respond positively to such methods assisting them in their typing, both subjectively in terms of user experience, and objectively in terms of their performance, including speed and accuracy.

As stated by Kukich (1992), error detection algorithms help with correcting "nonwords", helping with misspellings of single isolated words, and correcting misspellings of words with the current context in mind. These techniques use a combination of dictionary lookup, pattern matching and n-gram analysis to provide accurate and efficient results (Kukich 1992).These kinds of techniques have the advantage of being able to detect and correct a user's errors automatically and with no input from the user other than the text they have entered. This means that users are not slowed down in their typing as they do not have to take the time to manually correct their errors. Also somewhat tackled with this method is the problem mentioned above by Tinwala and MacKenzie (2010) of user's having to devote a significant amount of their attention to their screen. With automatic error correction, users are able to direct their attention to other tasks whilst entering text.

Alongside error correction algorithms and systems, prediction algorithms have also been utilised to speed up typing on mobile devices. These algorithms were utilised even during the prevalence of 12-key keypad devices (James and Reischel, 2001, Pavlovych and Stuerzlinger, 2003**,** Silfverberg, MacKenzie and Korhonen, 2000)**.** With touchscreen QWERTY keyboards, these prediction algorithms tend to strive to accomplish the same goal, as error correction methods, that is to help increase typing speed. By correctly anticipating the word a user wishes to type, the number of keystrokes required to compose a message or input is reduced overall, and can help fix mistakes, reorder sentences and increase the quality of the input (Vitoria and Abascal, 2005). As mentioned above for prediction algorithms for 12key keypads, QWERTY prediction algorithms have evolved to not just be able to predict words from a dictionary database, but also improves its suggestions depending on the context (Louis P. SLOTHOUBER, Eric H. Davis, Michael K. Young, Jeffrey W. JOHNSTON, 2008). Another advantage provided by these algorithms are the support they provide to people with disabilities (Vitoria and Abascal, 2005). In fact these methods were initially used to help people with disabilities, and were then adapted for widespread commercial use.

2.3. Tactile Feedback

Regardless of method attempted to improve text entry with mobile devices, common consensus seems to be that the lack of tactile feedback is they key limiting factor. This lack of tactile or physical feedback is in fact what has motivated the work towards error correction and prediction algorithms. While these are very creative solutions improving text entry on mobile devices, there are other areas to explore. Specifically the concept of actual physical or tactile feedback to users; there is still the potential of creating or designing novel innovative solutions or concepts that could help introduce actual physical feedback to text entry on mobile devices. Brewster, Chohan and Brown (2007) experimenting with the use of tactile feedback for mobile interactions found that “with tactile feedback users entered significantly more text, made fewer errors and corrected more of the errors they did make”. Hoggan, Brewster, and Johnston (2007) ran experiments to compare physical keyboards, non-tactile and tactile touchscreens. They found that tactile feedback on a touchscreen improved text entry, “bringing it close to the performance of a real physical keyboard, and concluded that the inclusion of tactile feedback on touchscreen devices would help in the betterment of touchscreen text entry. Rudchenko, Paek, and Badger (2011) when developing their game for improving user accuracy with touch screen keyboards state that “due to the lack of tactile feedback and generally small key sizes, users often produce typing errors”.

In the majority of work done towards tactile feedback development and research, it was found to not only improve user accuracy and typing speed in the vast majority of cases, but also to be very welcomed by users from studies, with significant positive feedback being received about tactile feedback methods. Dunlop and Taylor (2009) found that they were able to raise speeds and entry rates through the use of vibrational feedback, “raising speeds from 20wpm to 23wpm”. Brewster, Chohan, and Brown (2007) found that with their tactile feedback interactions users conveyed a “strong subjective feedback in favour of the tactile display”. As stated above, the main disadvantage QWERTY keyboards on smartphones compared to their 12 key predecessors is that there is no definite tactile feedback to be had from the input model. With this in mind, and the results of the works outlined above, it can be argued that finding a novel method of reintroducing this form of feedback could lead to significant increases in both user speed and accuracy when typing, as well as an overall greater of use and level of satisfaction from users from these methods compared to non- tactile methods. Already on a large majority on commercially available smartphones, a short vibrational “tap” is used to alert the user when they have hit a key when performing input, providing some sort of physical feedback.

A number of works seem to agree on the fact that the reason a lack of tactile feedback leads to loss of accuracy and typing speed is due to users having to repeatedly look at the screen to double check for errors and misspellings, rather than being notified of potential mistakes through some other means (Yatani and Truong 2009, Dunlop and Taylor 2009). This consensus helps to confirm the ideas outlined earlier by Tinwala and MacKenzie (2010)**.** Through of use of standard keyboards users can get automatic feedback when they are touching the edge of a key by feeling the edge of the key itself. With touchscreens, there is no such natural feedback, meaning that there is no intuitive way a user can deduce that they may have a made a mistake typing. While the use of predictive text entry and error check help mitigate this, Dunlop and Taylor (2009) argue that “as prediction gets better, users will drop the slow operation of checking the screen and will thus miss prediction errors and system feedback/suggestions". This furthers the case for vibrational feedback on devices for text entry being beneficial.

Yatani and Truong (2009) designed a user interface with tactile feedback named Semfeel through vibration motors attached to the backside of a mobile touch-screen device. This model was able to generate different patterns of vibration, and their testing showed that users were able to accurately distinguish between a high number of patterns. Dunlop and Taylor (2009) used vibrational feedback to alert users when word completion would be likely to help them. This was able to minimise their time pausing and looking at the device screen, increasing typing speeds in the progress.

There can be no doubt that tactile feedback combined with other methods such as prediction and error-detection algorithms have helped improve user accuracy and speed with text entry on mobile devices. Predictive text entry has helped users improve speed, and error detection has helped minimise the time users have to spend looking at the device screen in order to check for mistakes. As mentioned above however by Dunlop and Taylor (2009), the better these algorithms get, the less incentive users have to look at their screen, potentially resulting in more errors from these algorithms. Research into the use of vibrational feedback has shown this method can help alleviate balance out such weaknesses in the algorithms. By giving users an alternative method to look out for their own errors other than pausing typing to looking at their screen, speed of typing can be improved. The results of and Yatani and Truong's (2009) Semfeel has shown that through use of vibration motors attached to mobile devices, specific vibrational patterns can be detected by users. As such, it is worth investigating, and the aim of this project to find whether using patterned/directional vibrational feedback can help users improve their accuracy and speed by guiding their taps, letting them know when they have hit the edges of a key and giving similar feedback to what they would get from a physical keyboard. When typing on a physical keyboard, users will subconsciously adjust their typing to centre on the keys as they touch edges so subsequent typing is more accurate. By mimicking this feedback, users performing text entry on mobile devices would be able to their mistakes easier, as the edge vibrational feedback would let them know when they have not hit a key in the centre. They would also adjust their typing to be more centred on their keys, improving their accuracy in the long run.

2.4 Summary of Findings

Real computer keyboards give instant physical feedback that alerts users of their errors and drops in accuracy. This allows them to quickly rectify their errors as well as stay accurate when typing, thereby increasing their typing speed as well. This physical feedback also allows users to type without looking at their keyboard, essentially being able to look constantly at their input to check for any errors.

With the advent of mobile devices, initial 12key keypads, while having slower input rates than normal, still had 12 physical buttons that gave physical feedback when typing. This meant that, as with normal physical keyboards, it is possible to develop motor memory of the device and its keyboards, enabling for fairly straightforward text entry with limited mistakes. The fact that this motor memory and physical feedback enables looking at the screen to be able to read the input, so as to be immediately aware of any mistakes made greatly enhances the accuracy of text entry on any keyboard or keypad with physical buttons or features.

The biggest limiting factor with smartphone touchscreen keyboards is the lack of physical feedback from the buttons in the keyboard. This lack of physical feedback not only removes the “motor” effect of typing after getting used to a keyboard thereby decreasing typing accuracy, but also prevents the user from focussing their attention solely on their input. They are instead forced to divert the majority of their attention to their keyboard, periodically checking their input for errors, of which there could have been quite a few, making correcting them a large hurdle that decreases typing speed. To alleviate and counterbalance this deficiency, there has been substantial research into the field of autocorrection and word prediction algorithm, furthering on what has already been done for algorithms working on 12key keypads. These algorithms have progressed to the point where they are able to offer corrections and predictions that re context sensitive. These algorithms help to increase speed by allowing users to finish their words without having type it out completely, effectively decreasing the total amount of key presses required to type out and increasing users speed when typing. The correction algorithms help remove any errors during typing, increasing accuracy. However, the better these methods become, the less incentive users have to look at their input, most likely resulting in more errors in the long run.

Currently there has been work in the area of providing physical feedback through means other than the keyboard itself. The current most prominent one is through vibrations in smartphone devices. Through creating some physical feedback it has been noted in previous research that user's speed and accuracy can improve. They can be made aware of when they have pressed a button they did not mean to press, and can be made aware of when they have made a mistake (when error correction and word prediction could help them). Research done has proved that users can differentiate between different vibrational patterns on their devices, meaning that information can be conveyed to users through specific vibrational patterns and methods. Therefore there is merit in conducting research on how vibrational feedback methods can help improve text entry on mobile devices by designing and carrying out tests with a vibrational feedback method.

3. Methodology

3.1. Overview

It is important to remember the research questions and aims of this project. The questions ask how vibrational tap feedback on mobile devices can improve the speed and accuracy of text entry on these devices, as well as how users would respond subjectively to this vibrational feedback, including ease of use, and perceived usefulness. Therefore the main aim of this project is to investigate the effects and benefits of vibrational directional feedback on text entry on mobile devices.

The main objectives of this project are as follows:

- Design and assemble a prototype device capable of providing directional vibrational feedback for testing
- Design and implement a vibrational feedback model that will utilise this prototype device
- Analyse the usefulness of this model through user testing, taking into account speed and accuracy of user text entry with and without it, as well as user feedback about the performance and ease of use of the feedback model itself.

The prototype device for this project consisted of two actuators connected via stereo audio jack to a mobile device. This design would allow for vibrational feedback to be powered by unique stereo sounds played by the device they were connected to. This would play into the goal of implementing a vibrational feedback model, designing this model would revolve around mechanics of key presses and how they drive the unique vibrations coming from the actuators.

The vibrational feedback model revolved around having both the actuators attached to the left and right sides of the mobile device. This project aimed to assess the use of vibrational feedback for text entry. To do this, the feedback model was designed to be able to alert users when they have hit the edge of a key on a keyboard. As mentioned above in the literature review by Tinwala and MacKenzie (2010), one of the key limiting factors behind text entry on mobile devices, is that users have to split their attention between the keyboard they are typing on, and the text they are inputting. It is extremely difficult to type on mobile without looking at the keyboard being used. Therefore users are slowed down in their typing as they have to periodically check their screen for any errors or mistakes. Furthermore, since they cannot look at their screen constantly while typing, it is inevitable that they will miss some of the mistakes they make, which takes a significant toll on their accuracy, and the quality of their messages and content typed.

When using a normal, physical QWERTY keyboard, a user gets immediate feedback through their fingers if they have hit the edge of the key, and thus are immediately aware of the potential for a mistake. This project aimed to ascertain the effect of mimicking the feedback created in this scenario. The hypothesis was that if a user would be made aware of the effect that they had hit a key edge, then they would be aware of the fact that there is a higher chance of them having made a mistake while typing, prompting them to look at the screen to see their entry. With the user's having this prompt available to them, they should be able to spend more time focussing on their keyboard to type, and only divert their attention to their entry when the feedback model made them aware of a higher probability of error. This increased attention on the keyboard was expected to raise the users' typing speed, as well as their accuracy. Therefore, to achieve this, the individual actuators were placed to the left and right sides of the phone. Upon the press of the edge of a key, the corresponding actuator would vibrate (ie left side for left edge press and vice versa). This directional vibrational feedback was hoped to be able to provide the feedback, and through it the effects outlined above.

An application was built for the Android platform to facilitate a custom text entry method for experimentation purposes. This application contained a custom QWERTY keyboard purposefully built so as to be able to detect edge presses of the letter keys. To aid users' typing, word prediction, spell checking and correction functionality was integrated into the application. These functions served the purpose of helping mimic the environment commonplace with most Android phones when texting. It was through this application that sounds were played when key edges were hit in order to drive the actuators.

User testing was undertaken in order to gain qualitative and quantitative data for this project. Using the application, and the vibrational feedback, participants were tasked with entering a set of phrases, and tasked with entering each phrase as quickly and accurately as they could. Two vibrational feedback methods were used during testing. One set of phrases was done with a “mono” feedback model (both actuators vibrating on key press) as a control, and the other with “stereo” feedback (mono feedback in addition to individual actuators vibrating on key edge press). For each phrase set run, objective quantitative data was required in order to be able to assess how participant’s performed. This data included factors such as amount spelling mistakes, time to complete each phrase, backspace uses, edge presses etc. With this data, participants’ speed and accuracy when typing could be assessed. By comparing the two data sets (mono results vs stereo results) against each other, participants’ speed and accuracy with each vibrational feedback model, relative to other, could be assessed, and any improvements, or deteriorations in speed and accuracy could be measured and presented, giving adequate objective results for the purpose of this project’s aims and goals.

As mentioned above for this project’s research questions, it was also important for the raw results of text entry speed and accuracy to be accompanied by subjective feedback from participants regarding the feedback method. Using a series of questionnaires and survey forms, data regarding participant’s disposition to the vibrational feedback was gathered. General information deemed important to be gathered included how helpful the stereo feedback method was, how convenient and easy to use the system was, and whether participants felt the vibrational feedback made a significant difference in their text entry performance.

Using this subjective data in conjunction with objective data gathered during test, an assessment of the directional vibrational feedback model could be made. This data was used to evaluate the effects of the feedback method on typing speed and accuracy, as well as how easy the method is to pick up and get used to, and to utilise in a beneficial manner for text entry.

With the above requirements and general designs outlined, the practical work for the project could be undertaken.

3.2 Prototype

The main purpose of this project was to determine whether distinct directional vibrational feedback would have an impact on text entry speed and accuracy on mobile devices (specifically phones). In order to be able to test this, a physical prototype had to be designed that would enable this specific kind of feedback.

This prototype consisted of two actuators, specifically the Haptuator Mark II, obtained from tactilelabs.com (2016)[1], attached to the sides of a Motorola Moto G. These two actuators were small enough in order to be able to be attached to a mobile phone without making it too bulky or unwieldy to use. They also were able to be driven as a common loudspeaker. This meant that the software implementation of the feedback would consist of playing sounds on a mobile devices; having the actuators plugged into the phone's audio jack would translate these sounds as vibrations in the actuators themselves.

The actuators themselves had to have their ends soldered to two 3.5mm mono jacks. These jacks were then plugged into a mono to stereo 3.5mm convertor, which was then plugged into the phone. Sounds, obtained from freesound.com (2016), were you used to power the vibrations between the actuators. Initially it was found that lower frequency sounds triggered more intense and more easily noticeable vibrations in the actuators. However upon creating custom tones of 20Hz and 40Hz in Audacity (2016), it was found that the actuators responded quite poorly, with the vibrations being barely noticeable. The reason for this was determined to be the fact that single frequency tones and sounds don't use the entire body of the actuator, rather just a select section of it that responds to the frequency being played. This resulted in the weak vibrations observed. Sounds that covered a large variety of frequencies were found to be much more successful in creating vibrations that were powerful enough to be easily detected when holding the actuators. The sound of piano keys and chords being played were, in fact, particularly effective, and these sounds were used in the final application to drive the actuators. A single sound with a piano chord playing was found, and its default state with sound coming from both channels was used as the default "mono" sound, to be played during mono feedback, and on stereo feedback when the user hit the centre of the key.

[1] Tactilelabs.com. (2016). *Tactile Labs | Haptics accessories and kits for research and design*. [online] Available at: http://tactilelabs.com/ [Accessed 23 Aug. 2016].

To create the other two sounds, the audio from one of the channels (the quieter) one was deleted. This created the sound for one side, then the sound from the remaining channel was cut and pasted into the other channel. This ensured that the sounds for centre, left and right were identical in every aspect except for the channel they played from (and which actuators they triggered).

The placement of the actuators on the device was also important to be considered. They had to be placed in a position where a user holding a mobile phone naturally would be able to feel the physical feedback coming from both of them, and also be able to distinguish between the sides they were coming from. Originally the actuators were going to be placed on the middle of the back panel of the phone. This would enable a user to type using only one hand, by using the thumb of the hand holding the phone to input text. However it was found that positioning the actuators as such and holding the phone in this manner did not enable easy differentiation of the vibrations from the individual actuators. Due to their proximity to the hand's palm, all vibrations, whether from both actuators (mono feedback model), or from a single actuator (stereo feedback model) seemed almost identical and were hard to distinguish. This was most likely exacerbated by the fact that the large amount of material (electric tape, blue tac) required to secure the actuators to the position most likely dampened the vibrations.

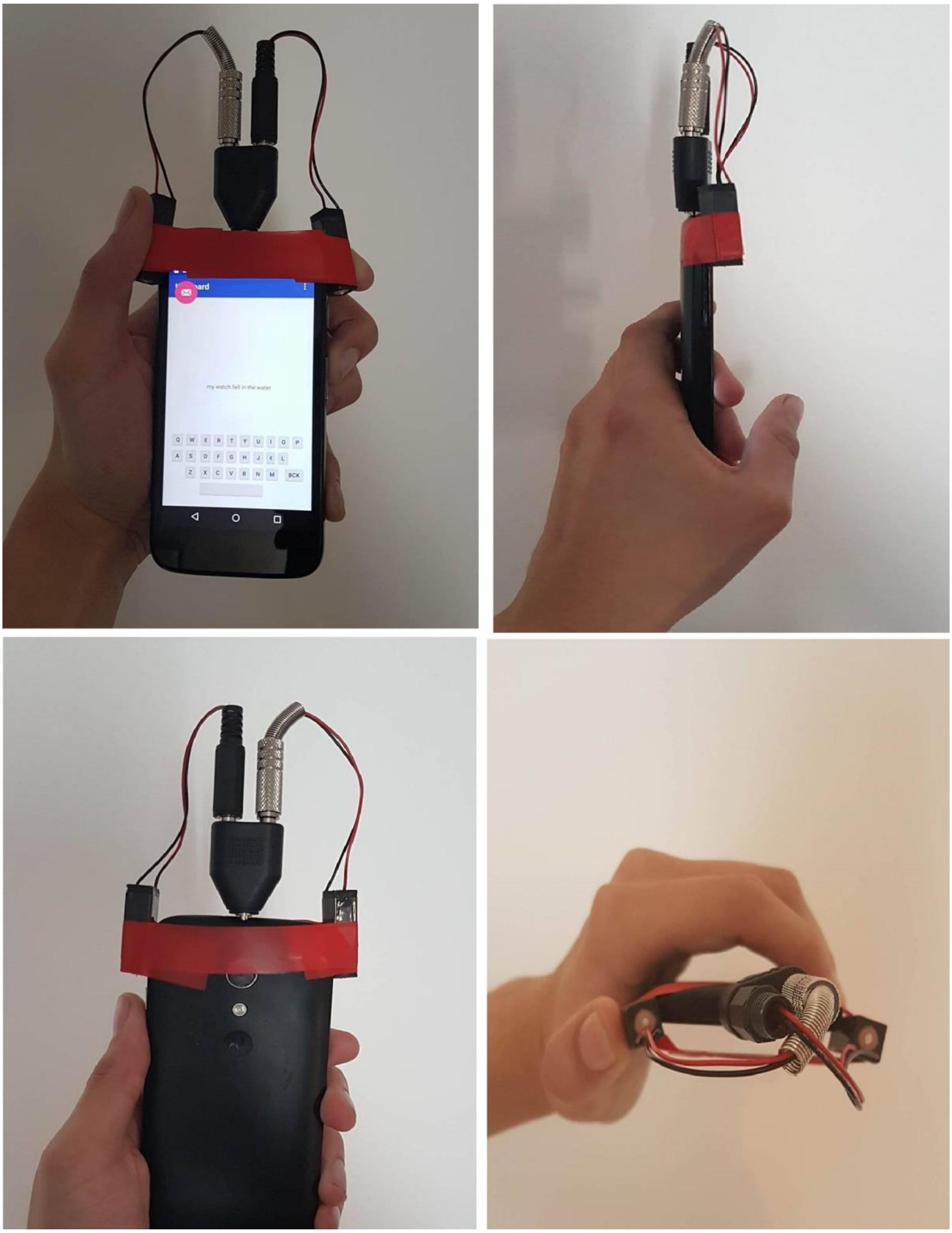

Figure 2 Prototype with final actuator placement

The final position settled on was putting the actuators on the top of the device, on the sides. The user would hold the device with hand, thumb resting on one actuator and index finger resting on the other. The index finger of the other hand would be used to input text on the keyboard. Due to the actuators resting on separate digits, it was hypothesised that it would be easier to differentiate between specific feedback from individual actuators in the stereo model. Less material was needed to secure the actuators in position, and this aspect definitely helped, as the vibrations were not dampened too much by supporting material.

3.3. Application

The application was developed for use and testing on the Android OS, and was developed using Android Studio. It consists of a custom QWERTY keyboard constructed for this project, an error correction system working based on Android's spell checker service, and methods for keeping track of the metrics used for this project, and writing them to a file on the mobile device for access.

On starting up the application on Android phones, an activity that contained a consent form for participants was displayed.

At the bottom of this consent form was a button labelled "I agree". This consent form is included as an appendix. Upon pressing the button, the main functionality of the app was made available to the participant.

3.3.1 Keyboard

The keyboard in the application developed for this project has its letters laid out in a classic QWERTY fashion. Users only have the capabilities of typing lowercase letters, as when comparing their input against the target input, the capitalisation of letters is not taken into account. All phrases presented to the user to type are completely lower case. A backspace button allows users to backtrack and delete their input one character at a time.

The keyboard itself consists of standard Android buttons all of the same size placed next to each other and laid out appropriately with regard to a QWERTY keyboard layout. All the letter buttons have the same custom listener developed attached to them. The backspace and space buttons have separate listeners attached to them.

3.3.2 Keyboard Touch Listener

The Keyboard Touch Listener class is a custom class developed for the letter keys that extends Android's TouchListener.

Upon touching a key (more specifically upon letting go of a letter key) this listener gets the X-position of the touch, relative to the letter key button (the left most edge of the key has X = 0). In the case of the stereo feedback model, if the press was on what has been defined as the left edge of the key, the appropriate sound is played that enables only the left actuator to vibrate. The opposite holds true for the defined right side of the button. Otherwise, if the key press was outside these definitions (i.e. the center), a sound is played that has both actuators vibrate. In the case of the mono feedback model, regardless of where the button is pressed, the mono, or “center” sound is played.

3.3.3 Suggestions and error corrections

Located above the keyboard are three boxes that contain suggestions for the current word being typed. The functionality behind these suggestion boxes are similar to the ones used on standard Android keyboards.

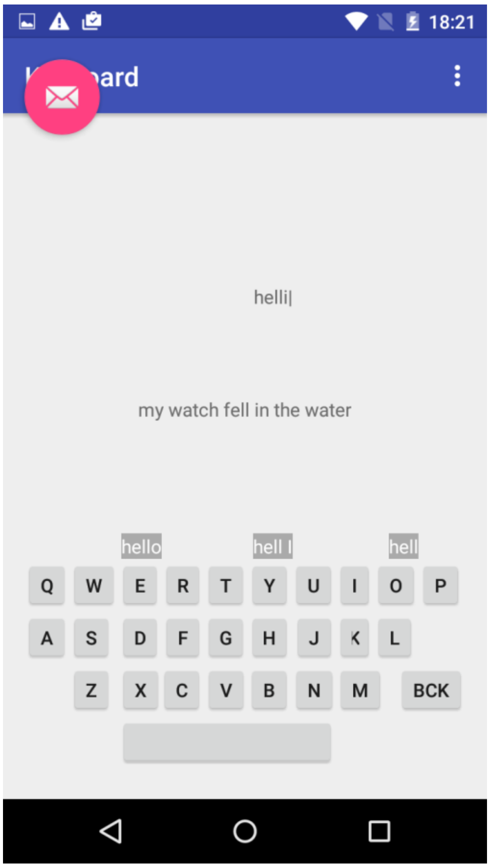

Figure 4 Application Keyboard

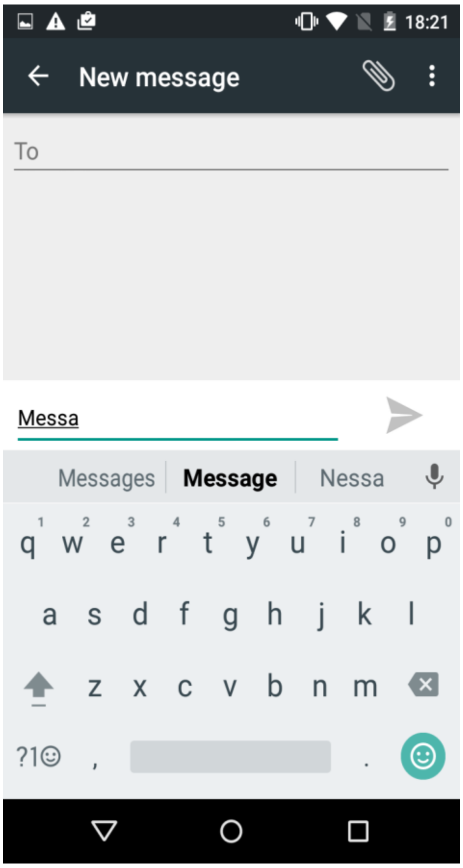

Figure 3 Android Keyboard

As mentioned above, these suggestions come from the use of Android's Spell Checker Service. The current word the user is typing in is used for this service, with the three most likely words being returned and stored in these suggestion boxes. The user has the option of tapping on any of these boxes, in which case the word suggested by them will replace the one they are currently typing and will allow them to type the next word. Use of these boxes in this manner has the potential of increasing user's speed when typing, as they have to press less characters to get their desired input.

The suggestion boxes also allow for an error detection system to be implemented that helps aid users' accuracy when typing. Using "Levenshtein Distance"(a method used to compare two Strings to determine how similar they are), the application decides whether or not to replace the currently typed word. Upon hitting the "space" button, if the word currently being entered is not in the dictionary, and the word's Levenshtein Distance is close enough to the primary suggestion from the Spell Service Checker, the word typed is replaced with this suggestion. This feature was meant to improve the overall accuracy of a user, by minimising the amount of words they entered that were erroneously spelled.

3.3.4 Phrases

Two sets of phrases were used for the user to type during experimentation and user testing. These phrases were provided by Mackenzie and Soukoreff (2003). These phrases are widely used for text entry studies. From the set of 500 phrases provided by the work, 100 were used to create the two phrase sets used in the application. The main advantages of these phrases were that they are "moderate in length, easy to remember, and representative of the target language" (Mackenzie and Soukoreff 2003). Another advantage of using these phrases is that they do not take into account punctuation and capitalisation, which meant that they worked with the keyboard designed for this application. The user of the application were displayed these phrases one by one, and were tasked with writing, using the provided keyboard the phrase displayed.

3.4. Experimentation Methodology

The experimentation was carried out by 9 participants. These were recruited through a mass circulated email to current university students, both undergraduate and postgraduate. Participants were also recruited through friend, colleagues and other personal acquaintances.

3.4.1 Participants' Application Tasks

As mentioned above, two sets of 50 phrases were provided for users to type in. Each user would do both phrase sets, and both feedback models, stereo and mono, in conjunction with these phrase sets. Participants were randomly assigned to conditions. This meant that the first feedback model and phrase a participant was used was randomised. Then the opposites of both variables were used in the second run. So for example, if a participant's first run of the experiment was with phrase set 1 and the stereo feedback model, their second would with phrase set 2 and mono feedback model. If their first was phrase set 1 with mono, then their second would be stereo with phrase set 2. Inevitably in text entry user studies, participants will be faster in the second question, as they have become more comfortable with their environment and the actions required of them for the experiment**.** This randomisation of mono vs stereo first served to counteract this learning effect in the final results, and to balance the condition. Similarly, participants did a different phrase set for each condition, which were also counterbalanced in order to ensure one phrase set isn't slower than the other due to the absence of this learning effect. This would produce a balanced experimental process, unaffected by this learning effect. This would also create a "Latin Square" design for the experimentation.

The actuators were held with electrical tape near the top of device, one attached to either side of the device. Users were instructed to, during testing, hold with their off-hand the phone near the top, with their thumb resting on one actuator, and their index finger resting on the other. During the stereo feedback model this was meant to make it easier for users to distinguish the individual actuators vibrations. All users were also instructed to type with the index finger of their main hand. This was in order to ensure that results obtained would be indicative of the difference between mono and stereo feedback. Testing conditions for all users would be constant, and the only variables present would be the mono and stereo feedback models.

The application was presented to participants, and they were given a few introductory phrases to type to get used to the keyboard, the suggestion boxes, prediction and correction systems, as well as the feedback coming from the actuators for the current feedback model. The intention of this short period of getting used to the feedback models was to ensure that the results of the experiments weren't skewed due to the fact that the participants were unfamiliar and uncomfortable with the input method (the custom keyboard, suggestion boxes, and correction), and the feedback method. After growing familiar with the environment for text entry, their resulting data and performance would be indicative of the feedback model themselves, rather than the participants' inexperience with the input method and feedback mechanics.

Participants were made aware of how to use the prediction boxes, as well as the mechanics and logic behind the correction method. They were made aware that the timing of the phrase started on first key hit for that phrase, and ended when they clicked the submit button go to the next phrase. As such users would not feel too much pressure going through the phrase set, and would not feel like they had to rush through the entire set. This assurance was to let them know that they had time to look at the entire phrase displayed on screen before typing it out. As mentioned above, one of the advantages of the phrases from Mackenzie and Soukoreff's (2003)phrase sets are that they were of moderate length, but still short enough to be easily memorisable. By being able to memorise the phrase that was to be typed, participants were provided with a reason to not have to look at the text they were inputting unless they believed they had made a mistake. This would help mitigate the problem outlined by Tinwala and MacKenzie (2003), in that users having to split their attention between the text input method and the results of their entry to periodically check for errors (errors that difficult and inconvenient to rectify on mobile devices) have their accuracy and speed drastically reduced while typing.

Unlike text entry in most mobile devices, the application developed for this project did not give users the ability to tap with their finger where they would like to type from and edit their entry from. The reason for this was to be able to assess how users fared in typing a phrase in one go. Giving them this functionality would allow for the results of the tests to be skewed in, as easier methods to correct mistakes would have minimised mistyping and errors in the entries. The aim of this project was to minimise the errors themselves while typing, and in order to keep track of them, it was deemed necessary that it would be difficult for users to hide these errors. To this end, participants were informed of limitations on using the backspace button. If an error they typed was instantly spotted, than they were free to use it. However if the error was only spotted after 3 or more characters, than they were requested to leave the error in. This was to have an indicative view in the final results of how many mistakes were made during typing and how often. Backspace usage alone would not have been sufficient data.

3.4.2 Metrics and stats tracked

The main question that the objective data gathered should be able to answer is, how has the use of the directional vibrational feedback model designed in this project affected text entry speed and accuracy for mobile devices. To this end, there were six metrics measured during user testing that would accomplish this goal.

- Time Taken – The time taken was measured independently for each phrase. The timing started for each phrase on the participants' first key press, and ended when they hit the submit button to bring up the next phrase. This was the primary metric used in determining the speed of the participants' entries for each phrase. Functionality such as the prediction and correction algorithms were included in the application not just to mimic the general text entry environment present in text entry on mobile phones(in particular Android phones), but also to allow users to achiever faster times on their entries on their phrases.

- Edge presses – The purpose of this metric was to investigate if the stereo feedback model helped users minimise the amount of times they pressed the edge of a key. With the mono feedback participants would most likely notice the amount of time they might have hit a letter on the keyboard off the centre. In the case of the stereo vibration however, it was possible that the repeated feedback from the method over the 50 phrases could improve the participants' overall accuracy, and reduce the amount of time the edge of a key was pressed.

- Backspace Count – This was the amount of time a participants used the backspace key in a phrase to correct a mistake. As mentioned above, users were instructed to only use the backspace key if they noticed the mistake they wished to correct almost as soon as they made it. If all it took to correct their current error was one or two backspaces, than they were encouraged to correct it. However if more than three or four characters (included spaces) had been entered since they error they entered, they were instructed to leave it as it was. The purpose of the stereo feedback model was to notify users of a potential mistake when they hit the side of a key, and prompt them to look at their entry to check for mistakes. Ideally, a backspace would occur when this feedback let the user know they have made a mistake, and would be indicative of how many times the feedback model helped them spot and correct their error. Using a large amount of backspaces to correct an error made earlier could hide a mistake that the feedback model potentially failed to alert the user of, whilst still drastically increasing the amount of backspaces the participant used for that phrase.

- Levenshtein Distance – This metric is used to measure the difference between the phrase from the phrase set (from now referred to as the target string) and the corresponding participant input (from now referred to as input string). The greater the value of Levenshtein Difference, the more different the two phrases.
  The implementation for calculating this variable was obtained from wikibooks (2016) entry on the algorithm, under the Java section. The implementation consists of a two-dimensional array to store the distances between the two phrases being compared. With the algorithm, deletion (a character missing), insertion (an extra character inserted), and replacement (a wrong character) were all giving a weighting of one. Therefore, for each one of these cases occurring, the Levenshtein Distance for the individual phrase entered increased by one.
  The purpose of this metric was to determine how close in similarity the input phrase the participant had entered was to the target phrase. By comparing the Levenshtein distances of stereo input phrases against those of mono input, an improvement in the quality of input could be identified. This would be evidence of an improvement of user accuracy when typing.

- Auto Corrects: As already mentioned, the three suggestion boxes in the application were populated by the top three suggestions from the Android Spell Checker Service.

The correction algorithm corrected words the user was entering if certain conditions were met. These conditions were:

- If the word the user entered was not in the dictionary
- If the Levenschtein difference between the current word being entered and the primary of the three suggestions from the Spell Checker Service was not more than 1

The auto correct functionality would only occur when the user hit space on the word being tested for an auto correct. Ideally autocorrects would occur less frequently in the stereo than in the mono feedback models.

- Manual Corrections: These three suggestion boxes could also be clicked. Doing so would replace the current word being typed with the word contained in the clicked suggestion boxes. The purpose of this was to help users type by reducing the total amount of characters they would have to type in order to type out the phrases presented to them. This would serve to decrease the time taken metric. This metric was tracked in order to see if any differences in the time taken metric between stereo and mono were due to the different feedback models, or an increase in manual correction usage. As mentioned above, the learning effect of participants could have caused them to get used to manual corrections and rely on them more in order to type first. Therefore this metric was tracked in order to counterbalance this learning effect.

3.5. Questionnaires and Surveys

Alongside the information about user speed and accuracy when typing, data needed to be gathered about the participants' general attitude and sentiment towards the feedback model. The question required to be answered was how users would respond subjectively to this vibrational feedback, including ease of use, and perceived usefulness. A feedback model that performed poorly with respect to objective data but that was well received by users could be could incite further research into that area of text entry for mobile devices. Likewise, positive objective data could be countered by poor user reception. In order to gather this subjective data, a series of questionnaires and survey forms were used.

An introductory questionnaire was given to participants before they started any testing. This questionnaire, relatively short compared to the other ones presented to participants, asked general questions to gain information about participants' race, gender, and level of experience using keyboards and text entry methods on smart phone devices. This experience could be deduced by two questions, which asked how many messages a day on average they sent, as well as how long they had been using text messaging methods (qwerty keyboards, tapping letters, drawing words, eg swype).

Two NASA Task Load Index (TLX) forms were used. One was given to participants to fill out after completing their first phrase set, and the second after completing the other phrase set with the other feedback model. By comparing these two against each other, the effect each feedback model had on the participants could be assessed. Information regarding how mentally demanding the task was, how the participants felt they performed doing their tasks, and the level of frustration incited by the tasks of the experiment. Through these questions, the psychological response users had to the feedback methods could be assessed. Since the questions posed by this project regarding the directional vibrational feedback model included how "distracting" it was when typing, and how easy it was to get acclimatised to and use alongside the input method, the answers provided in the TLX form were deemed an apt way to deal with the these questions.

The Task Load Index, as describe by NASA (2016) is a subjective workload assessment tool which uses "multi-dimensional rating procedure that derives an overall workload score based on a weighted average of ratings on six subscales". Questions asked include mental demand, physical demand, the participant's performance, effort and frustration. This form could assess the participant's subjective view of the workload from the tasks relative to each feedback model tested. These forms were handed out after each phrase set, and as such after each feedback model. Comparing the results of the two side by side would lead to an apparent distinction between the two in terms of task workload, and the feedback models' usefulness. The ideal results for this comparison would have been a smaller rating for workload and exertion with a smaller level of frustration, and higher rating for the user's own performance when using the stereo feedback model. However, an identical rating for exertion, workload, and frustration combined with a higher rating of the participant's own performance would still be enough evidence for an apparent subjective improvement over the mono feedback model with the stereo vibrational feedback.

The final questionnaire was given to the participants after they had finished the second NASA TLX. This questionnaire was designed to pose more specific questions about the difference between the two feedback methods than were asked in the TLX forms. Questions included asking participants for their preference between the two feedback methods (and if they even noticed a difference between the two), what exact impacts they felt the stereo method had on their typing when compared the mono feedback method. Through these open-ended questions where users could elaborate on their opinions on the success and effects of the feedback models, specific positives and faults of each of the models could be obtained. When combined with the objective data gained through the participants' typing performance, the feedback and opinions obtained could be used to develop a conclusion on the benefits of the directional vibrational feedback model developed for this project.

The above outlined methodology was undertaken in order to answer the research questions for this project. Specifically, these questions were:

- What effect will a directional vibrational feedback method have on the speed and accuracy of users when typing on mobile phone devices?
- How useful and beneficial will users find the feedback from the method?

With the data obtained from participants' input when undertaking typing tests on the application built, the first question can be answered. Through metrics such as words per minute, edge presses, and how close participants' inputs were to the target phrases, their speed and accuracy with the mono feedback model could be compared against the data of the stereo feedback model.

The second question was attempted to be answered with the questionnaires and TLX forms filled out by participants. Through this, participants made clear about the usefulness of the stereo feedback model, as well as its convenience and ease of use.

An analysis of this data combined allowed for conclusions about the success of the feedback model for improving text entry on mobile devices. Information gathered would include potential improvements to the vibrational feedback mode, suggested by participants or otherwise, as well as suggestions for future work in the area of research this project concerns itself with.

4. Analysis of Results

The quantitative data gathered from testing was written to two files for each user, one for mono testing, and one for stereo testing. These files were run through a parser so that the data could be analysed to gain conclusions from the testing.

In the first run of the parser, an average of each metric for each phrase across all users was taken, for both mono and stereo tests. For example, an average was taken for the levenshtein distance across all users when entering phrase 1 (in both mono and stereo). This was repeated for all phrases, and all metrics, for both feedback models. This was done to have an overview of the differences between the individual values for the metrics produced by the mono and stereo feedback models. By looking at these overviews for each metric, substantial differences between the mono and stereo feedback model averages could be identified, which would prompt further investigation into the specifics of this metric. These user averages where calculated using phrases 20 – 50, rather than 1 – 50, in order to be able to discount any learning effect that would inappropriately skew the information gathered.

It was also necessary that the results of the questionnaires (subjective and qualitative data) be taken into account alongside the analysis of the quantitative data. The results provided by the metrics may not have been in line with what can be gathered from the opinions provided by participants.

The participants had an average age of 27 years (with a standard deviation of 5.36). Of the nine participants, three were female. The participants reported on the amount of texts they send a day, and the majority reported to send 20 or more a day, with the next highest majority being 10 or more a day. The length of time these participants reported to having owned a smartphone and used text entry on touchscreen averaged to about 5 years.

The metrics are presented as averages per phrase across all users, for stereo and mono, as well as an average of the metric per user across phrases 20 - 50 in stereo and mono. The reason for not taking the first 19 phrases into account is so the results are unaffected by any learning curve from participants getting used to their test environment in the initial few phrases.

4.1 Levenshtein Distance

The first metric to be looked at was, Levenshtein Distance, was believed to be the most straightforward one to get started with, and to get some ideas about the success of the vibrational feedback model with. The expectations behind this metric were that a lower Levenshtein Distance would be an obvious indicator of overall accuracy a user achieved while typing. In line with many lab studies in text entry the error rate was very low in both conditions. A preferable result would be a lower overall Levenshtein Distance for phrases input with the stereo feedback model, as opposed to the mono feedback model, as this would signify an improvement in typing accuracy obtained through use of the feedback model.

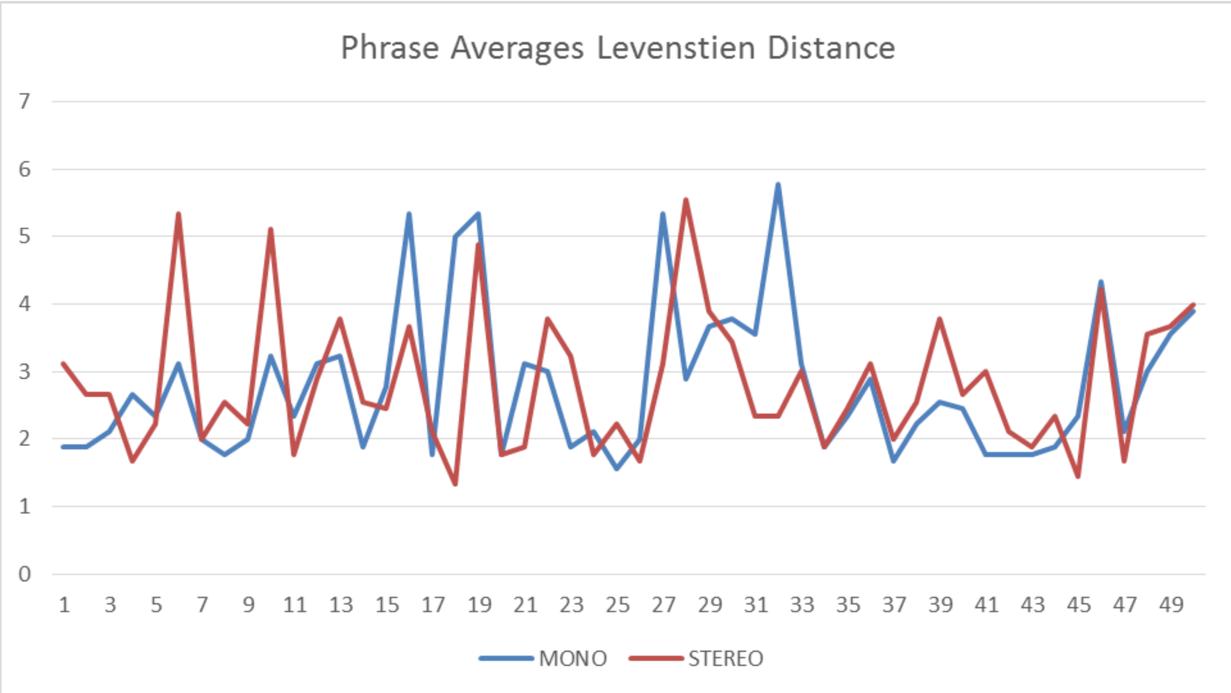

Figure 5 Levenshtein Distance average for phrases across all users

The graph above shows the average Levenshtein distance for each phrase across all nine users. From the graph, it can be noted that there were no substantial and easily identifiable improvements provided by the stereo method. There are cases where there is a lower average Levenshtein distance for the phrase, however these see in the minority, and any improvements seem rather small. It seems more common that mono and stereo's Levenshtein distances were close to being the same, or the mono having a lower Levenshtein distance.

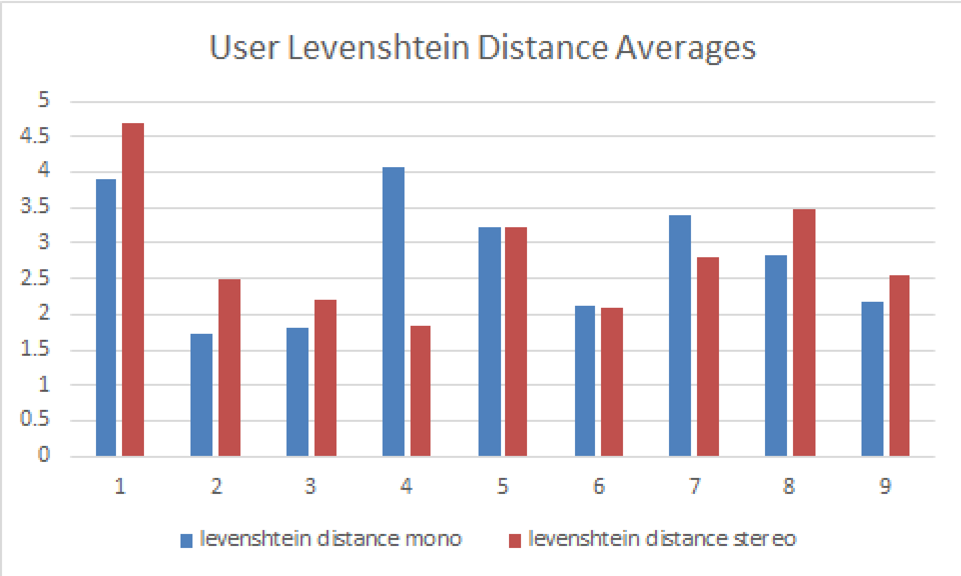

Figure 6 Levenshtein Distance User Averages

The graph above showing each user's averages for Levenshtein distance for each test helps to confirm this hypothesis. An equal Levenshtein distance, or greater with stereo is more common than any improvement with the stereo method.

t-Test: Paired Two Sample for Means		
	Mono	*stereo*
Mean	2.807407	2.818519
Variance	0.793272	0.773086
Pearson Correlation	0.419779	
Hypothesized Mean Difference	0	
df	8	
t Stat	-0.03496	
P(T<=t) two-tail	0.972965	

Table 1 Paired T Tests with user averages for Levenshtein Distance

The paired T test above shows a very low difference between the overall means of the data pictured in the graph showing user averages. The t stat is greatly smaller than that of the t critical (two-tail) value, meaning there is no significant difference between the means. A p- value of 0.05 was used, and the two-tail value is much greater than this, further confirming that there is no substantial difference between the two means,

4.2 Words per minute

A metric measured during testing was time taken for each phrase. This was then converted to a words per minute through the formula:

$$Words\ per\ mintue = (\frac{\left(\frac{phrase\ length\ in\ characters}{5}\right)}{\left(\frac{time\ in\ seconds}{60}\right)}$$

This metric was the most straightforward metric for analysing the speed of participants' text entry. In the hopes of proving that the stereo feedback model provided an improvement in typing speed over the mono feedback model, higher average times for the phrases entered was deemed a preferential observation of this metric.

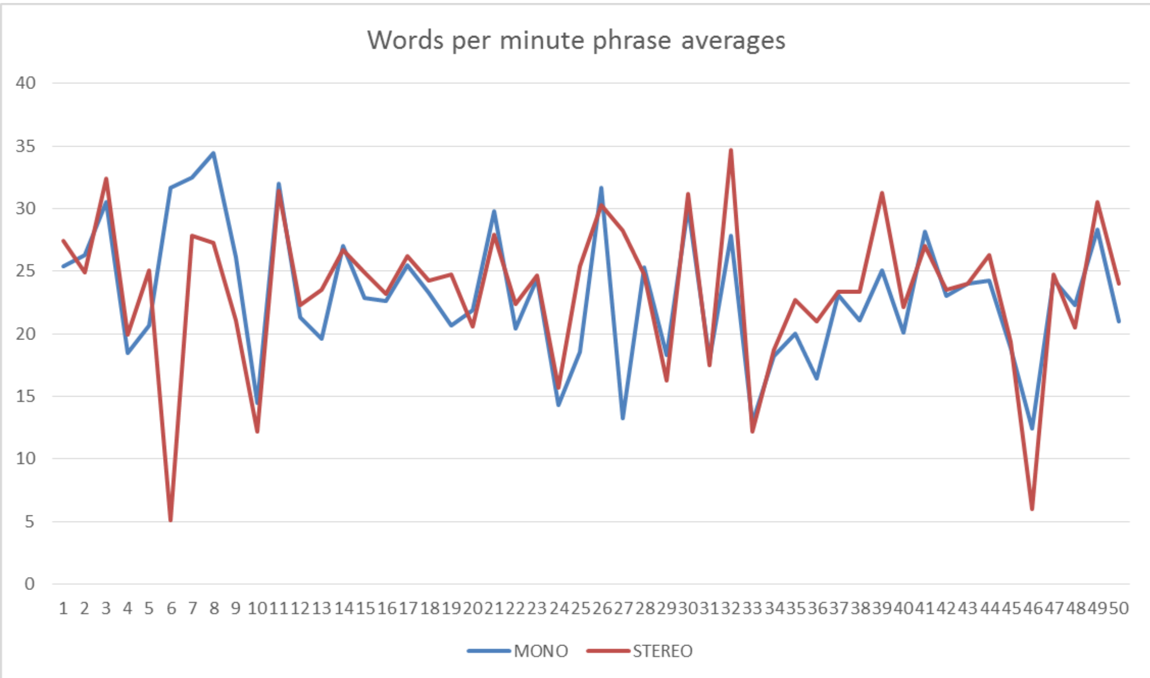

Figure 7 Words per minute average for phrases across all users

At first glance, the stereo feedback mode does not seem to provide an easily quantifiable improvement over the mono model. The graph suggests that overall, the words per minute averages were near identical. However on closer inspection it can noted that with the stereo feedback model, a greater words per minute count is consistent. Whilst there are cases where mono's words per minute are equal to or greater than stereo's, these results are less common than a greater value for the stereo feedback model.

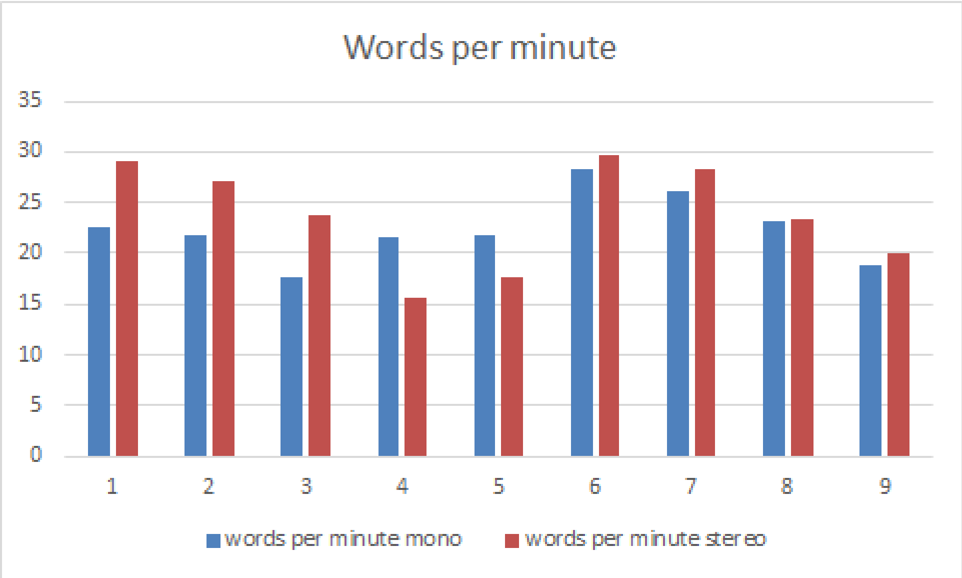

Figure 8 Words per minute user averages

With the user averages in the graph above, this theory can be reinforced. With the exception of two of the users, the mono feedback model never seemed to outperform the stereo feedback model in terms of the input speed.

t-Test: Paired Two Sample for Means		
	Mono	*stereo*
Mean	22.42357	23.83143
Variance	11.06982	26.64132
Pearson Correlation	0.558642	
df	8	
t Stat	-0.98132	
P(T<=t) two-tail	0.355174	

Table 2 Paired T tests with user averages for words per minute

However upon performing a paired t test, doubts on these claims are created. A two-tail p-value greater than 0.05 confirms, rather than denies the null hypothesis that there is no substantial difference between the typing speeds of the two methods. This is furthered by the fact that the two means are near identical.

4.3 Backspaces

Backspaces were measured in order to be able to see if the feedback model was successful at alerting the users when they had made a mistake. By notifying the user that they had pressed the edge of a key, they would check their input to see if an error had occurred, and if so correct it through the use of the backspace key. This is the reason that participants were instructed not to correct any mistakes they had made that they didn't notice immediately, as the larger amount of backspaces required to fix those would have polluted the results. The purpose of this metric, in conjunction with the Levenshtein distance metric, was to determine if a reduced Levenshtein distance of stereo inputs with respect to mono inputs was due to users correcting their mistakes more often as they were able to be made more aware of them.

The backspaces were tracked per phrase. With the above in mind, the preferred results would have been a higher backspace count in conjunction with a lower Levenshtein distance. Seeing as in the section above it appeared that there was little or no improvement in Levenshtein distance during stereo tests, this ideal result is unlikely. However, looking at the difference between backspace counts in mono and stereo testing can still provide some information as to the effects of stereo vibrational feedback versus mono. It is worth mentioning that due to the "lab" environment, users were very likely making much less mistakes than they would in everyday use. If users' performance were recorded whilst they were walking or doing something else while typing, performance would have been more representative of the actual effects from experimentation.

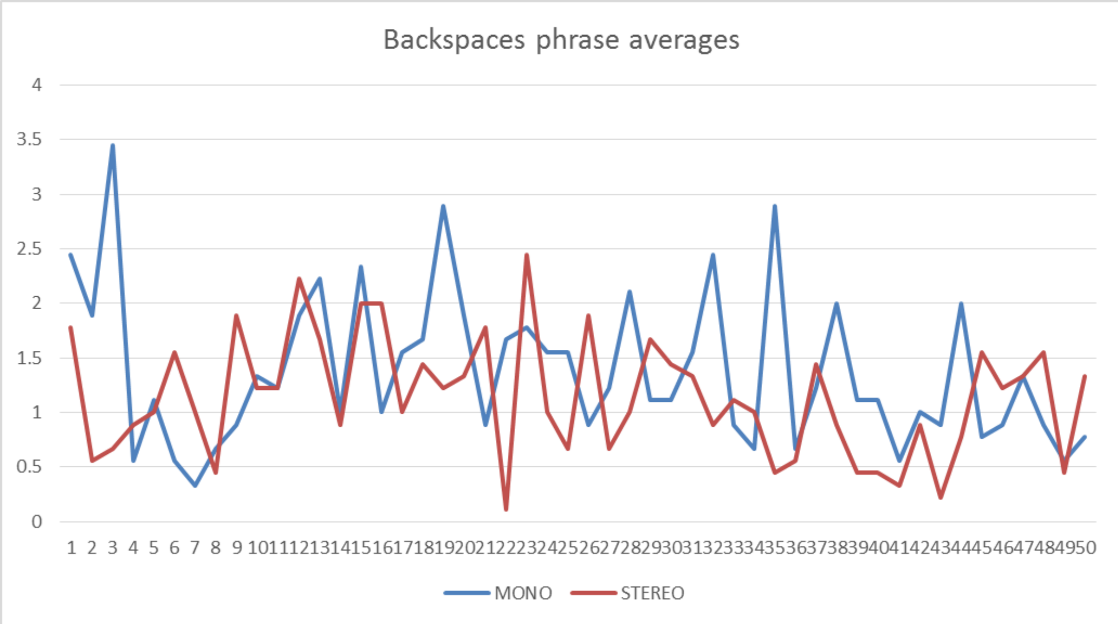

Figure 9 Backspaces average for phrases across all users

The above graph appears to suggest that there was a larger amount of backspaces used on average during mono testing than with stereo. While this was not an expected result, one possible explanation could be that with stereo feedback users were made aware of how often they pressed edges, and could have been adjusting their grip and typing to be more centred on the keys. This is in fact the very same effect physical keyboards are able to give their users through physical feedback. Looking again at the above graph, it can be noticed that in later phrases during the stereo feedback, the averages for the backspace counts were got lower, in contrast to the increase in backspaces for the mono method in later phrases. This observation could support the hypothesis that users adjusted their typing as they were made more aware of how often they hit edges. With this information in mind, observing the difference in edges pressed could prove conducive to showing that the stereo feedback did in fact have some positive effect on feedback.

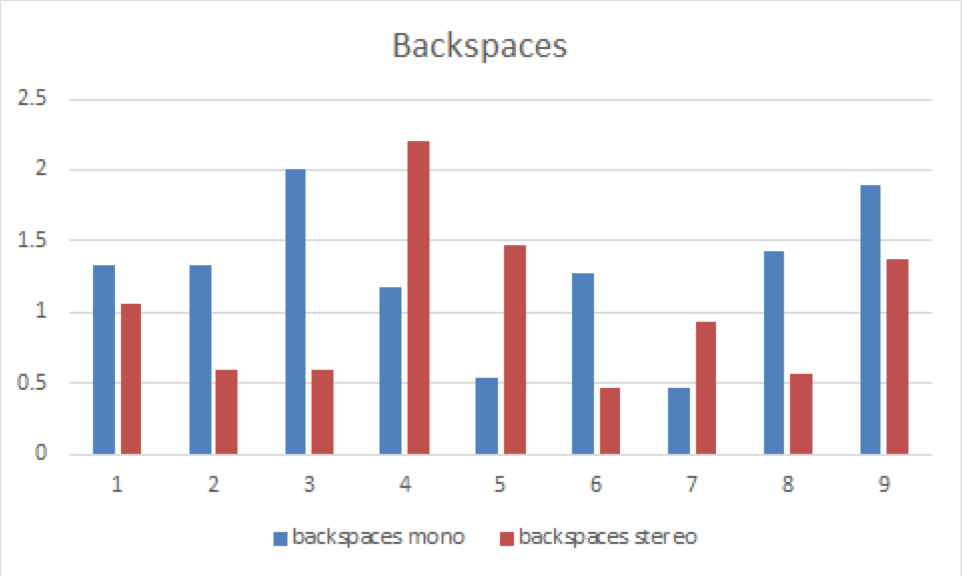

Figure 10 Backspaces user averages

t-Test: Paired Two Sample for Means		
	mono	*stereo*
Mean	1.27037	1.02963
Variance	0.270401	0.323179
df	8	
t Stat	0.842452	
P(T<=t) two-tail	0.424003	

Table 3 Paired T tests with backspace user averages

However, as with the other metrics, the p-value from paired t tests is lower than the alpha value of 0.05, once again confirming the null hypothesis that there is no significant difference between the two feedback models in relation to backspace counts.

4.4 Edges

As mentioned above in the backspaces section, any potential changes in backspace counts could also be accompanied by changes in edge presses. Whilst a paired t test of user averages for backspace count proved that there was no substantial difference between the stereo and mono method, it would still be appropriate to check if there were any substantial difference between edge presses. A decrease in edge presses in the stereo method would be indicative of the feedback model being able to let users become more aware of how often they press the phone edges, and how often these edge presses can be indicative of mistakes made during typing. This knowledge would allow them to adjust their typing in order to be able to more often hit the centre of the keys, thereby increasing their accuracy in the long run.

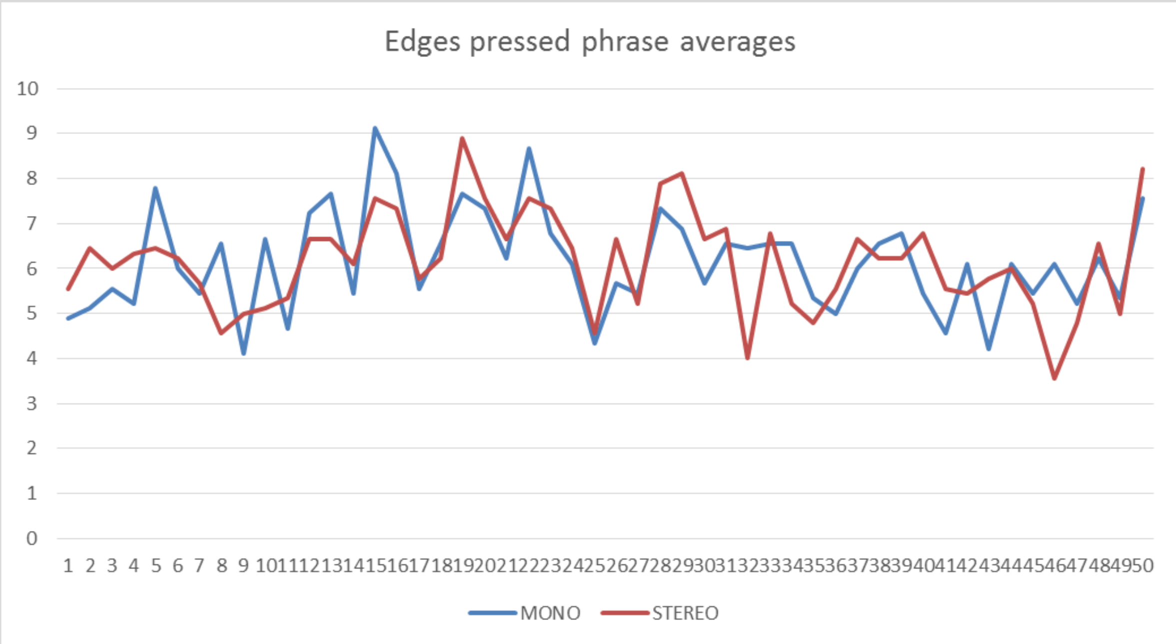

Figure 11 Edge presses average for phrases across all users

In the diagram above, both the mono and stereo and methods seem to follow the same curve. An initial increase in edge presses (most likely due to the learning factor of typing on a new keyboard), with an eventual dip in the amount of edges pressed. These two curves seem to be fairly similar, showing little difference except in a few phrases.

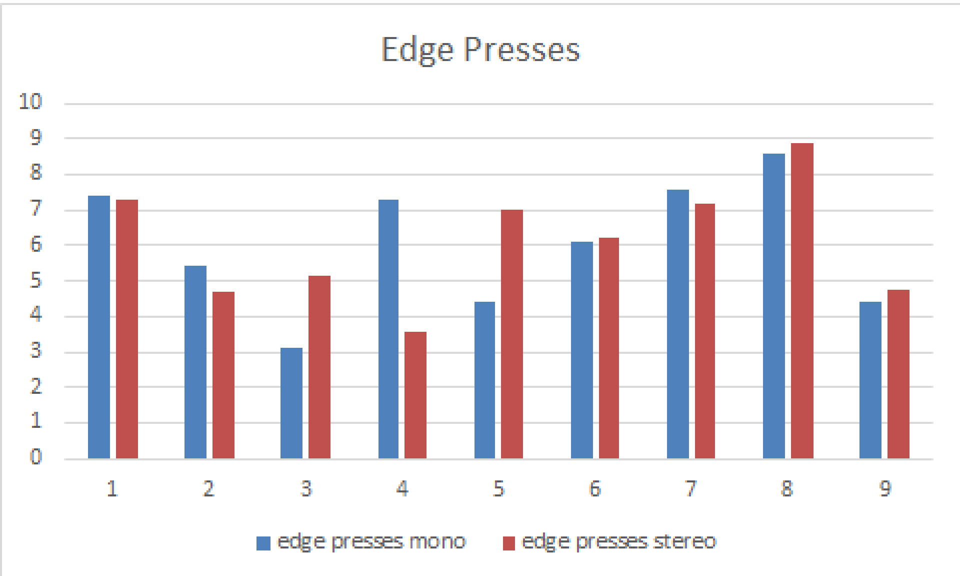

Figure 12 Edge presses user averages

This lack of difference of edge presses between the two feedback methods is reinforced with the above graph of user averages. The differences noted are often relatively small and seem to be inconsequential in the grand scheme. There is no consistent improvement apparent from the stereo feedback method.

t-Test: Paired Two Sample for Means		
	mono	*stereo*
Mean	6.040741	6.077778
Variance	3.266883	2.801667
Pearson Correlation	0.477555	
df	8	
t Stat	-0.06232	
P(T<=t) two-tail	0.951839	

Table 4 Paired T tests with user averages

The paired t tests on the data from the graph of user averages only serve to confirm this lack of notable difference. As with the other metrics looked at thus far, the null hypothesis is confirmed, and on significant improvement is noted as a result of the stereo vibrational feedback method.

4.5 Auto corrections

From the data obtained in user testing, it was noted that the restrictions imposed on autocorrection were strict enough it stop it from ever triggering. This resulted in an average of 0 auto corrections for each phrase, regardless of the feedback method. Therefore, this metric is not used in the evaluation for the feedback method.

4.6. Manual Corrections

The purpose of manual corrections was to help users type by reducing the total amount of characters they would have to type in order to type out the phrases presented to them reducing the time taken for each phrase. Manual corrections were tracked in order to see if any differences in the time taken metric between stereo and mono were due to the different feedback models, or an increase in manual correction usage. Any learning effect of participants could have caused them to get used to manual corrections and rely on them more in order to type first, so the metric was tracked in order to counterbalance this learning effect.

In the analysis of the words per minute metric it was determined that there was no notable or easily distinguishable difference between the two methods. Normally, this would warrant checking the manual corrections metric for the reasons stated above. However considering that for all other metrics, including edge presses, backspaces, and Levenshtein distance analysis yielded similar conclusions, there are no irregularities that can get in the way of explaining why there was no difference between the time taken results for the two feedback models. These irregularities would warrant an analysis of the manual correction data, however given that there was no difference in all other metric analyses, it was clear that manual corrections did not have an effect on user performance or speed. However for consistencies sake, the data from manual corrections will still be included below.

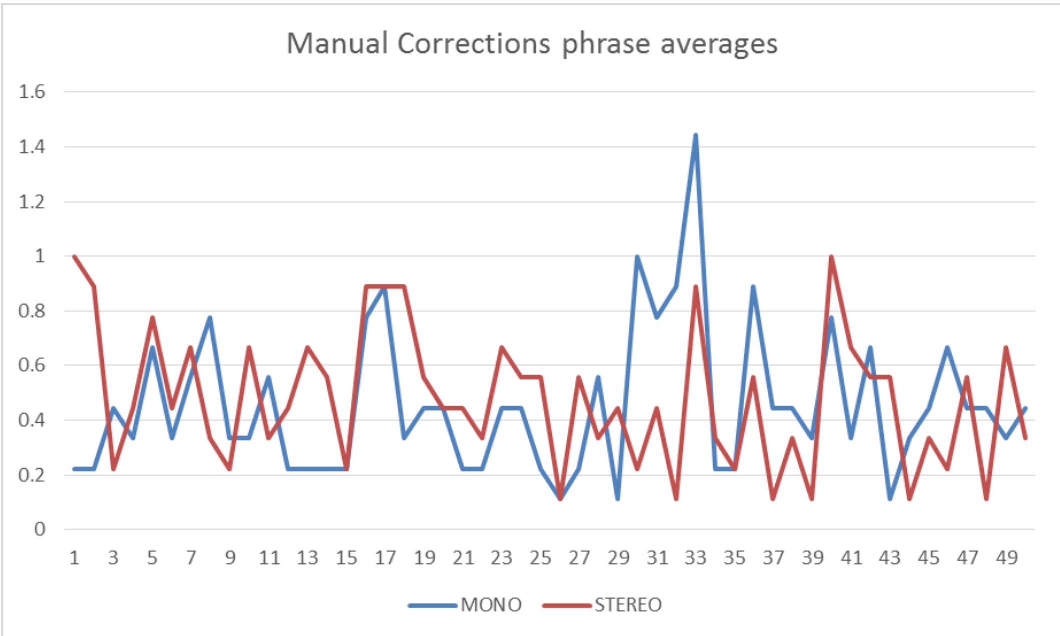

Figure 13 Manual Corrections average for phrases across all users

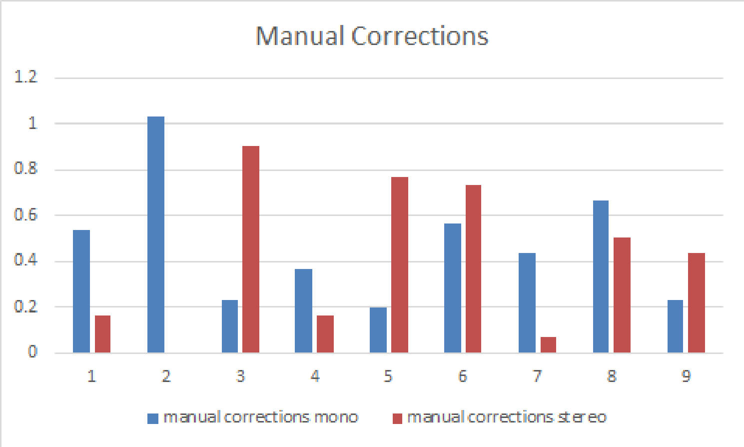

Figure 14 Manual Corrections user averages

t-Test: Paired Two Sample for Means		
	mono	*stereo*
Mean	0.474074	0.414815
Variance	0.070772	0.110586
Pearson Correlation	-0.53361	
df	8	
t Stat	0.338534	
P(T<=t) two-tail	0.743674	

Table 5 Paired T Tests with manual corrections

4.7 Questionnaires and Qualitative Data

The qualitative data from the participants comes from the information they provided in the two TLX sheets, and the final questionnaire provided at the end of testing.

The answers given in the TLX forms provided indicated that at some level, users were able to distinguish some differences between the stereo and mono feedback models. The most notable differences were in the physical and mental demand of the tasks, with most users implying that the use of the stereo feedback model was more demanding than that of mono feedback. Perceived performance varied greatly between users and feedback methods, with some users reporting an increase in performance with stereo feedback. The majority however, seemed to be reports of no difference, or increased performance in mono. These filled out TLX forms will be attached as an appendix.

To read the TLX data, the participants' ratings for each topic were taken as a number between 1 and 21. The lower this number the better the result and the more preferable the feedback model. The graph below contains the averages across all users for all the topics in the TLX form, in both Mono and Stereo.

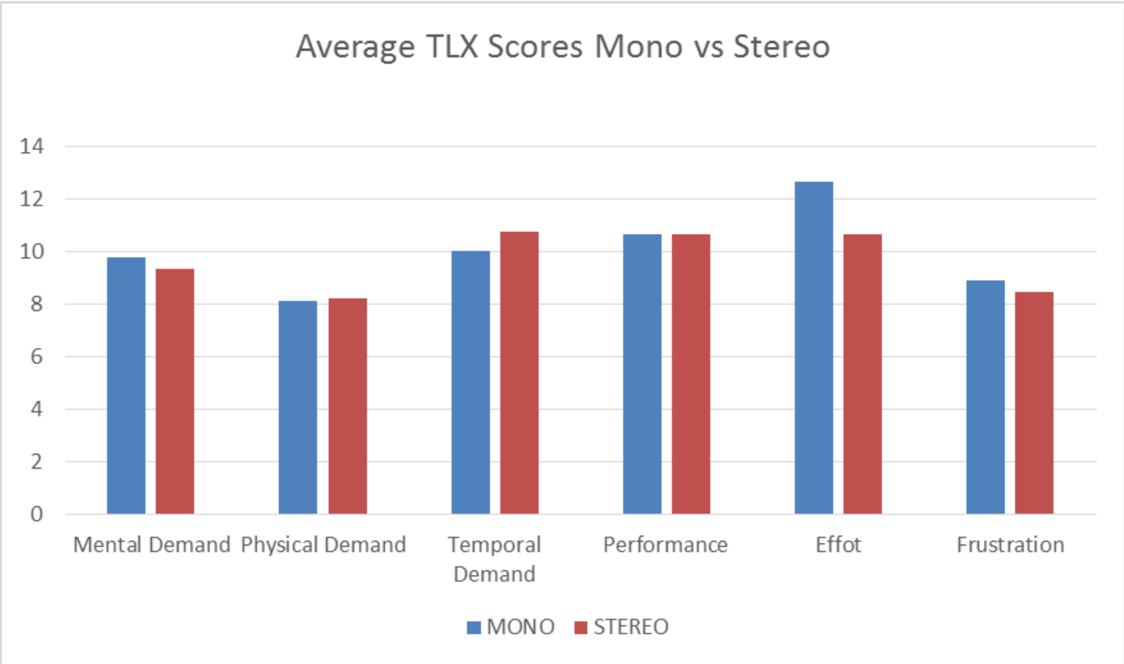

Figure 15 Average Tlx scores for each category across all users

The graph below shows participants' tlx scores for all 6 topics summed up.

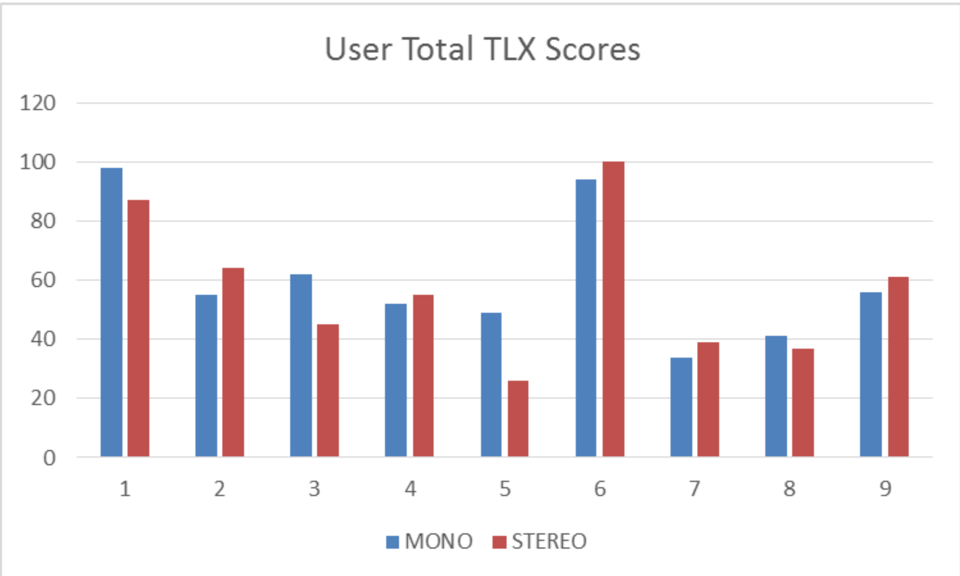

Figure 16 User total summed tlx scores

The table below shows the paired p test for these total scores.

t-Test: Paired Two Sample for Means		
	MONO	*STEREO*
Mean	60.11111111	57.11111111
Variance	483.3611111	580.8611111
Pearson Correlation	0.880372999	
df	8	
t Stat	0.785584405	
P(T<=t) two-tail	0.454746135	

Table 6 Paired P test for user summed scores of TLX forms

The graph of averages seems to show a lower overall score for stereo feedback, making it the more favourable of the feedback models. However the graph of the scores summed shows more similar results between the stereo and mono feedback models. The p test above performed on these sums shows a p value greater than 0.05, showing as with all other metrics, that the null hypothesis is true and there is no substantial difference between the feedback models in this regard. At this point it can be easy to assess the effects of having a relatively small amount of participants for a study, as while the data seems to favour conclusion that boasts a difference between the two elements being tested, statistical analysis of the data shows a conclusion that stats there is none.

The answers given in the questionnaire provided at the end of testing contained more specific with the feedback from users.

One of the main risks of failure for the stereo feedback model was that users may not in fact feel any difference at all between the stereo feedback and mono feedback. Potential reasons for this included that the vibrations might not be strong enough, or distinct enough to individual identified in the case of the stereo feedback. Of the nine participants, eight reported that they were able to notice a difference between the two feedback models.

Of the eight participants that were able to notice a difference between the mono and stereo feedback models, seven reported that stereo had any sort of impact on their typing. Five of the eight reported that they preferred the stereo feedback model. Reasons given for this preference included a perceived increase in how often the participants noticed their mistakes immediately after making them. This increase was attributed by the participants to the directional feedback of the stereo model. However it was reported that it was not the vibration itself that was the cause of this positive feedback, but rather the sound causing it. When playing sounds through the actuators, the sounds could still be heard quite faintly. Users were able to differentiate between the “mono” sound playing through both channels and the stereo sounds that would play through either exclusively the left or right channel. This is particularly true of participants that did their first phrase set with the mono feedback model, as they had grown accustomed to, through the learning a factor, a “normal” sound, this being the mono feedback.

Of the three participants that preferred mono feedback, one gave their reason to be that the stereo model was too distracting. They felt the feedback in fact slowed them down, as they had to stop after each piece of directional feedback to assess the difference between this piece of feedback and the general mono feedback given when hitting a centre of a key. Another user noted that they felt the simple mono feedback that they let them know whether they had hit a key or not was sufficient enough tactile feedback for them to be able to type accurately and quickly enough.

The final question on the questionnaire was how easy the stereo feedback method was to get used (if they were able to notice a difference between it and the mono feedback model). This was presented as a scale from one to seven, with one being the hardest, and 7 being the easiest. The average was 5, with a standard deviation of 1.65.The majority were numbers including 5 and above, indicating the learning curve of the stereo feedback model was not too high.

5 System Documentation

5.1 ConsentActivity

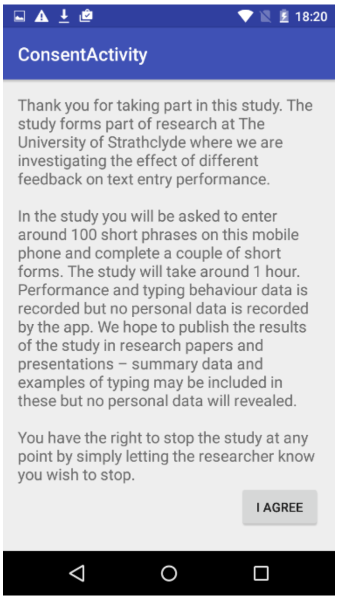

Figure 17 Consent Activity

The "ConsentActivity" is the startup activity for the android app. It consists solely of a Text View object that has the text contained in the consent form.

At the button of this text is an Android button labelled "I Agree", than when clicked sends an intent to start the MainActivity.

This ConsentActivity ensures that users have read the consent form and understand roughly what the experimentation will entail, how long it will take, and that they have the right to leave at any point. By having this as the startup activity, it is ensured that users cannot enter the MainActivity for testing until they have read and hit "I agree" to the consent form.

5.2 File Writer

The file writer is used to create and write to the testing device's memory the file containing the data used in storing the participant's performance data when testing. Passed in the data in the form of a String from MainActivity, and the new file, the file has the data written to it via the use of an instance of the Print Writer class and its print function. When the data is written, the Print Writer object is closed.

One of the main difficulties of writing the data to the android device was manually accessing files written to the Android device's memory. Unless the device was rooted, or was written to the SD card, the file with the data written to it could not be accessed through PC file explorers, or through the phone itself. Therefore it was necessary to install a third party, ES File explorer, in order to gain access to the folder the files were being written to. To access these files outwith the mobile device, they were shared via email to an address where they would be downloaded onto a computer.

5.3 Keyboard Touch Listener

The Keyboard Touch Listener class was an implementation of Android's On Touch Listener class. The purpose behind this class was to build a custom on touch listener that would handle all the responses necessary when a user presses one of the letter keys on the custom keyboard built in the application.

The main onTouch function handles what happens when the button is pressed, and more specifically, determines these responses based on where exactly on the button the press was detected. The Main Activity contains an Android MediaPlayer object. By passing the main activity into the KeyBoard Touch Listener class by reference when they are created, the listener can decide what feedback to give by playing sounds through this mediaplayer. Checking if the currently selected feedback model is mono or stereo, the mediaplayer plays the appropriate sound. If in mono, then the mono sound that makes both actuators vibrate is played, regardless of where on the key the press is located. Otherwise if in stereo, if the press is located on the designated left or right side of the button, then the corresponding sound is played that will make only the actuator on that side vibrate (i.e. a key press on the left edge of a key will make a sound play a stereo sound that only has sound coming through the left channel, and vice versa). A press on the centre of the key plays the same mono sound present on all key presses when mono feedback is enabled.

```
float x = event.getX();
//if the left edge of the key is pressed
    if (x < 20) {
        if(mainActivity.getStereo()) {
            mainActivity.playNewSound(R.raw.soundleft);
        }
        else {
            mainActivity.playNewSound(R.raw.soundcentre);
        }
        mainActivity.incrementEdgesPressed();
    } else if (x > 55) { //if the right edge of the key is pressed
        if(mainActivity.getStereo()) {
            mainActivity.playNewSound(R.raw.soundright);
        }
        else {
            mainActivity.playNewSound(R.raw.soundcentre);

        }
        mainActivity.incrementEdgesPressed();
    } else { //centre of the key has been pressed
        mainActivity.playNewSound(R.raw.soundcentre);

    }
```

Figure 18 Keyboard Touch Listener code snippet

The listener also has a reference to the Text View that contains the user's input, located in the main activity. When created, the custom touch listener has a String called "letter" initialised, corresponding to what key on the QWERTY keyboard the listener is attached to. When they key is pressed, the text view is updated with the letter appended to the end.

It is also worth noting here that the "cursor" present in the text view is not in fact a movable cursor that users can move to alter where their input will be inserted. It is simply the "|" character that is just removed from the string in the text view every time it changes, and then appended back at the end of it.

5.4 Levenshtein

The Levenshtein class contains the Java class obtained from wikibooks' page on Levenshtein implementations.[2] It contains the main function that is used in the main activity at the end of a user's input to compare their entry to their target phrase.

---

[2] En.wikibooks.org. (2016). *Algorithm Implementation/Strings/Levenshtein distance - Wikibooks, open books for an open world*. [online] Available at: https://en.wikibooks.org/wiki/Algorithm_Implementation/Strings/Levenshtein_distance [Accessed 23 Aug. 2016].

5.5 Phrases

The Phrases class is simply one that contains two string arrays, each containing all the fifty phrases of a phrase set. They were stored in a separate class in order to stop the main activity from becoming too cluttered with the definition of two large string arrays at the top.

5.6 Suggestion Touch Listener

The suggestion touch listener, similarly to the keyboard touch listener, is an implementation of Android's On Touch Listener. This listener was used with the text views that contained the suggestions from the Android Spell Checker Service implemented for this application.

```
@Override
public boolean onTouch(View v, MotionEvent event) {

    //replace the word being typed currently with the word in the suggestion box
    String text = view.getText().toString();
    text = text.replace("|", "");
    String words[] = text.split(" ");
    String lastWord = words[words.length-1];
    String suggestionText = suggestion.getText().toString();
    String newText = text.replace(lastWord, suggestionText + " " + "|");
    view.setText(newText);
    mainActivity.incrementManualCorrection();

    return false;
}
```

Figure 19 Suggestion Touch Listener code snippet

The purpose of this touch listener was to change the word the user is currently typing for the one selected from the suggestion boxes. To this end, and again similar to what was implemented in the keyboard touch listener, the listener has a reference to the main activity, and the text view that contains the user input. When pressed, the listener gets the text of the input, and removes the cursor "|" character. By separating the input by spaces, a string array of the words entered is obtained. The final element of this array is the last word of the string, or more accurately, the word currently being typed. This word is replaced with the word in clicked suggestion box, adding a space and the "|" symbol so the user can immediately continue typing the next word. Touching a suggestion box also calls the main activity's "incrementManualCorrection" function, that increments a counter used in the main activity to keep track of manual corrections as a part of the data being tracked.

5.7 Main Activity

The main activity class is where all the functionality in the app is presented to the user. It contains the keyboard, the suggestion from the spell checker service, and the phrase text views.

The buttons of the keyboard were placed manually in the main activity's content xml. Buttons of the same size were placed in the appropriate position in order to mimic a QWERTY keyboard. These buttons were then instantiated in the main activity and added to an array list of buttons, ordered from top left to bottom right of a QWERTY keyboard (q, w, e, r, t, y, u, I,…etc , b, n, m). A string array containing individual letters of the alphabet was ordered in this same manner. This made the string and list to be of the same length. Looping through the list of buttons, they were all added with a new keyboard touch listener, with its letter being that specified in the string array. As such, when each of these buttons were pressed, the text view in the main activity was updated with the letter from that key by the key's keyboard touch listener.

The backspace button (and its listener) had similar functionality to the keyboard touch listener. Upon clicking the button, the text in the text view was obtained as a string. The "|" character was removed, and then the final character in the string was removed as well. Then the "|" character was again re-added. This action also incremented "backspaceCount", an integer used to keep track of the amount of backspaces used per phrase as a part of the data collection for experimentation.

The suggestion boxes consisted of three text views placed above the keyboard for easy access. These suggestion boxes were stored in an array of textviews, and had any modifications made to them collectively through this array, including setting their listeners. The suggestions obtained from the spelling check service were placed in these textviews, in order of relevance (highest to lowest) according to the service, from left to right.

The space button's listener did more than just add a space to the text view. It was also what triggered the auto-corrections to occur. Upon pressing space, the word just entered was obtained in a similar manner to the suggestion touch listener (i.e. splitting the string by spaces after removing the "|", and getting the last element of the resulting array). This word was compared against the primary suggestion obtained from the spelling check service, (i.e. the left most suggestion box). The levenshtein between these two words was computed. If the word entered by the user was not in the dictionary, and the two words had a levenshtein distance no greater than 1, than the entered word was automatically corrected to this primary suggestion. A space and the "|" key were readded in order to enable the user to continue with their entry.

A function called "go" was used to make calls to the spell checker service and get recommendations for the words being entered. Every time the input text view was updated, this function would be called. The spell checker session would get the word currently being entered (with the same method of extracting from the text view as outlined above), and provide three suggestions for spell checking using the onGetSuggestions method. These suggestions would then be inserted into the suggestion boxes above the keyboard.

Initially development was being undertaken on a Samsung Galaxy S7, and experimentation was expected to be done on this device as well. However it was found necessary that the Locale of the spell checker session be set to English simply to not cause the application to crash. Furthermore, despite best efforts, the spell checker service was not providing any suggestions when prompted. The reason was eventually deduced to be the fact that the spell checker service is disabled on Samsung devices. It was in fact this reason that influenced the change of the testing and development device for the prototype developed to be a Motorola Moto G.

At the top of the activity was a menu that allowed control of the settings og the application for experimentation. These setting included setting which phrase set to use, as well as whether to use mono or stereo feedback. Both of these were achieved using appropriate checkboxes in the menu.

The Floating Action Button in the main activity was used by the user to submit the entry for the current phrase and proceed to the next one. Upon pressing this button, all the metrics tracked were recorded and added to the string that would contain all the data written to the file on the device.

- The Levenshtein distance would be calculated immediately between the phrase entered and the target phrase.
- Backspaces, manual corrections, edge presses and autocorrects were being tracked through the use of separate integers that incremented every time one of the above occurred.
- Time taken was tracked by getting a the current time the first time the input text view change, then getting another at the point the FAB button was pressed, then getting the difference between the two. This information was stored as seconds

If there were still more phrases to go in the phrase set, then the FAB button would simply add the current data to the string containing the participant's data, reset all the counters, time taken, and levenshtein distance, and set the next phrase to display. However if the final one of the set had been entered, then the data for the final phrase would be added to the string, and a new file would be created. This file would follow the following naming structure: Date, stereo or mono feedback, which phrase set was being used. Then, this file and the string of data would be passed to the file writer where the data would be written to the device.

6. Conclusions and Recommendations

This project undertook an investigation of the effects of directional vibrational feedback on text entry in mobile phone devices. Directional vibration to mimic the physical feedback available from keys on traditional keyboards was used, to let users know when they have hit the edge of a key, that they have may have made a mistake, and to allow them to, over the course of their typing, adjust their input to be able to more often hit the centre of the key they are aiming for.

The main questions of this project ask how vibrational directional feedback on mobile devices can improve the speed and accuracy of text entry on these devices, as well as how users would respond subjectively to this vibrational feedback. From these questions, the main aim devised for this project was to investigate the effects and benefits of vibrational directional feedback on text entry on mobile devices.

6.1 Analysis of work done

Once again, the main objectives of this project are as follows:

- Design and assemble a prototype device capable of providing directional vibrational feedback for testing
- Design and implement a vibrational feedback model that will utilise this prototype device
- Analyse the usefulness of this model through user testing, taking into account speed and accuracy of user text entry with and without it, as well as user feedback about the performance and ease of use of the feedback model itself.

To meet this objectives, the prototype device for testing was constructed, as was the Android application in order to be able to implement the vibrational feedback model. Experimentation was done to gain quantitative data with which to assert the affect the feedback model had on user's performance when performing text entry. Surveys and forms given to the participants throughout testing served as the method of gathering the subjective data necessary to determine how the participants responded to the feedback method. An overview and analysis of how well these objectives were tackled with the work done follows.

6.1.1 Prototype

The prototype assembled worked well to test out the two different feedback models. It fulfilled its main purpose, which was to help get an understanding of whether mimicking the physical feedback of key edges would be noticeable by users, and bare any impact on their typing.

The largest question concerning the prototype was to do with the actuators on the device, or more specifically the ideal location to put them. The actuators were placed in the position were users were most likely to notice the different vibrations coming from them during experiments. However due to this, users were also instructed to hold the device in a very specific way (hand near the top of the device, thumb resting on one side, index finger on the other). Whilst this condition ensured that participants had the best possible chance of feeling vibrations, it may have been an uncomfortable or unusual position for them. In future work, a method of providing vibrations that wouldn't force users to hold the device in a specific way and was more flexible would be preferable.

Whilst more complex vibrational patterns could have been explored, similar to what was done for **Semfeel**, considering the technology available to the prototype, the feedback model of left vibrations for left edge presses and right vibrations for right edge presses was deemed the most appropriate method to implement. It was also considered to be the most straightforward and easy to grasp concept for participants.

6.1.2 Application and Implementation

Participants reported that for the most part the application developed for testing was straightforward and easy to use. One point of complaint was that the application sometimes tended to erase words in about four of the phrases (phrases that began with the word "I"), leading to increased frustration for users when typing. This error was down to the autocorrection occasionally firing off in circumstances where it shouldn't have. Due to the fact that this error was only reported after the experimental process was about four or five participants in, it was deemed inappropriate to fix it for the remaining participants, as that would produce unfair and inconsistent results towards the end of the experimentation.

One other slight issue that users reported when performing their tasks for experimentation was that there was a slight delay for vibrations when hitting a key. Most users that noticed this delay stated that at times it was a bit distracting, however not enough to be able to detract from their performance in any meaningful way. This reported “laggyness” was determined to come from the applications “MediaPlayer” object. By having to load the appropriate individual sound used for feedback each time before playing it, this slight delay became apparent the more a user used the application. In future, some other more preferable method might be used, that could preload all the sounds and not have to do it again to play them, to avoid having such a delay when playing the sounds.

Overall, users reported that the application was very straightforward and easy to use. This might owe to the fact that the keyboard in the application was very simple, in that users did not have to worry about punctuation or capitalisation. However no issues were reported with the keyboard, manual corrections, or any other aspect of the application. Participants who had any sort of trouble using the keyboard reported that it was due to their increasing reliance on word drawing methods on phones, such as swype, rather than any inherent problem with the keyboard in the application itself.

6.1.3 Experimentation

The experimentation done was planned in such a way in order to ensure the fairest representation of the results of the directional vibrational feedback method, balancing the benefits of a controlled usability lab environment with task representativeness. Steps undertaken, such as the randomisation of what phrase set and feedback model to start with was included in order to counterbalance the learning factor that might be present (and influence results) if all users started with the same phrase set and feedback method.

The data gathered during testing was very useful in determining the objective effect of the stereo feedback model during text entry. All the metrics gathered had a purpose behind them, and how could they be interpreted in order to determine the effects of the stereo feedback model in relation to the mono feedback model was clearly defined before the start of experimentation. These metrics allowed for definite conclusions about how the stereo feedback model impacted text entry on the prototype mobile devices.

The same holds true with the surveys provided to participants. The subjective feedback given by participants in the forms was plentiful and varied enough in order to be able to answer the question of how the users responded subjectively to the feedback model. The questions posed in these surveys asked the exact questions necessary in order to determine what was aimed to be known about the feedback model by the end of this project.

The one major drawback to the experimentation, and such to the project overall, was the relatively small amount of participants that undertook the experimentation. While the data gathered contained noticeable trends, and lent itself to forming conclusions about the data and the project's work, a larger amount of participants would have been preferable. With only nine participants, the objective data available was in fact limited in how much it could show, and what conclusions could be drawn from it. Whilst it was assumed that certain participants would improve their typing with the feedback model, and others wouldn't, nine participants was, in hindsight, not enough to be able to determine any particular trends or favoured results between those gathered from stereo and mono feedback. The results, obtained from averaging the data gathered from specific phrases, and from the individual users, each with both feedback models, were far too similar to be conducive to drawing any conclusions on the positive effects of the stereo feedback model. Likewise, concerning subjective data, it was assumed that some users would respond positively to the stereo feedback method, and others wouldn't. Whilst there was a majority favouring one stereo feedback method over another, a larger amount of participants might have lent itself well to a more clear and overwhelming majority, and substantially more feedback for the stereo model.

6.2 Results of Data

6.2.1 Quantitative Data

After looking at the words per minute, Levenshtein Distance, manual corrections, backspace and edge press counts metrics, no conclusive evidence could be found that there was any overall positive effect the stereo feedback model had on the typing performance of participants. When looking at the graphs in the analysis section of this report, some patterns could be made out that would suggest favourable results leaning towards the stereo feedback model, the paired T tests showed that there was no noticeable difference between the two sets of results. Thereby the conclusion that could be reached is that this sort of tactile feedback had no beneficial impact on text entry.

This conclusion that stereo feedback model had little or no effect on the participants' performance directly conflicts with what the conclusions that can be derived from the qualitative data obtained from participants (detailed below). An overall (but not overwhelming) majority positive reaction to the stereo feedback model suggests that participants not only found the feedback method to use, but also felt it had a beneficial influence on the way they typed.

This conflict, and lack of difference in effects from the two feedback methods is most likely caused by the small sample size of participants used in testing. Preferably, more participants would have been able to provide more data. With a larger amount of data, more definite conclusions could have been reached on the effect of the stereo feedback model. The fact that the data currently suggests that there was no discernible difference between the two, and yet the subjective data suggests otherwise, allows the conclusion that it would be worthwhile repeating this experiment, but with a larger sample size in order to gain more definitive a more reliable idea as to the positive influence the stereo feedback model could have.

6.2.2 Qualitative Data

As mentioned above, the majority of users responded positively to the feedback model. They reported that they found the feedback was able to reliably make them aware of their mistakes, and it also helped their accuracy in the long run as they slowly adjusted their typing to be able to more consistently hit the centre of the keys.

Though these users did report that they perceived the feedback to have a positive effect on their typing speed and accuracy, this was more from the difference in the actual sounds played through the actuators, which were still prevalent alongside the vibrations, than from the vibrations themselves. Participants reported that they found it easier to discern between the mono sounds and the other two stereo sounds, than their vibrations. So while there was a noticed difference between the two methods, it should be noted that it more often than not did not come from the vibrations themselves.

Some users reported that the stereo feedback method was a bit too distracting, and in fact slowed them down in the long run because of this. This distracting factor could have come from the sound lag issue outlined above. As stated on Android's site (2016), "anything that takes more than a tenth of a second to respond in the UI will cause a noticeable pause and will give the user the impression that your application is slow." Using an alternative method to MediaPlayer, such as the SoundPool class, might have more favourable results in future.

Overall, almost all users were able to discern a difference between the two feedback methods. And of these, the majority reported that believed the stereo feedback method had a positive influence in their typing. When to describe this positive influence, users stated that they were made away of mistakes from the stereo feedback sounds, and that they felt their accuracy improved in the long run, as they adjusted their typing to more often hit the centre of the keys. They reported that they were not aware of how often the hit keys off centre until the stereo feedback pointed it out to them. This in fact describes both of the goals of this stereo feedback method, namely to prompt users to look at their input only when there is a high likelihood of them having made a mistake, and to improve their accuracy over time by having them adjust their typing to less often make mistakes form hitting the sides of keys.

It is also worth noting that an overwhelming majority of users, even those who did not feel the stereo feedback model benefitted them, reported that the feedback was easy to get accustomed to and use.

As mentioned above, this response from participants directly conflicts with the conclusions that can be gathered from the results of the objective data analysis. The same point will be made again, in that this conflict can be attributed to the small amount of participants. Having more data to work with could have helped draw out more reliable conclusions about the feedback models.

6.3 Recommendations for Future Work

It is clear from the results outlined above that this project would have benefited greatly from a larger amount of participants. Repeating the methodology and experimentation with a larger sample size of users would be greatly beneficial in gathering more data about participants’ performances when typing, as well as their responses to the usefulness of feedback model, and the quality of their own text entries.

More prominent vibrations would help users greatly in being able to correctly identify the feedback they are receiving. The sounds used were the ones that produced the most noticeable feedback. If there any sounds that would have any improvement on the range of the vibrations, then these would obviously would preferred. As noticed while testing out different sounds, the actuators seemed to respond most to piano notes and chords being played, as this got a response along their full body, causing more prominent and noticeable vibrations.

What could also help the vibrations to stand out is to remove the delay between pressing a key and the feedback received that some users reported. This delay was most likely the reason that some users reported the feedback to be distracting. It is imperative that the feedback given to be fast and for its meaning to be easily discernible, and removing this slight delay would help both of these goals to be accomplished.

It could also be noteworthy to investigate the use of sound as a method of feedback alongside the vibrations. While vibrations are most likely more beneficial overall then sound, as users may be in areas or situations when typing were sound be inappropriate, the fact that users were reacted more to the sound coming from the actuators could mean that it is an avenue of research worth exploring. Experimentation with audio feedback would also most likely be easier to carry out, making any results and feedback from this work also be conducive to helping research into the use of vibrational feedback as well.

6.4 Final Conclusions

This project successfully developed a prototype text entry method using stereo tactile feedback to give feedback on users typing in the centre, left or right edge of keys. The project investigated different patterns of feedback and selected strong, short vibrations that would attempt to mimic the instant feedback provided when typing on a physical keyboard. A 9 participant formal usability experiment was conducted that showed that users could perceive the difference between mono feedback and stereo feedback (however they were not able to reliably tell the difference between left and right). This was mainly through audio difference with tactile stereo effect being very muted given the actuators were placed on the exterior of the device, and the dampening effect of the materials required to secure them. There is some evidence that the stereo effect helped with speed of entry and while no significant result was found in this study it would be worth repeating the study with increased number of users. While the studies showed a majority of users typed slightly faster with stereo feedback there were no significant differences and opportunities for studies are discussed. These results indicate that it would be worth repeating the study with increased number of users.

Overall, this project was able to tackle the questions posed by it. The work included constructing a physical prototype to successively emulate how the vibrations for the feedback model would feel, building an application that housed a keyboard allowing specific feedback based on where a key was pressed, and conducting experimentation to obtain quantitative and qualitative results. Through this work, research conducive to furthering the field of tactile feedback for text entry on mobile devices was done, and results were gathered that would help future work. This area of research has worthwhile benefits on text entry, and hopefully the work outlined by this report will help drive further work into it.

References

Audacityteam.org. (2016). *Audacity®*. [online] Available at: http://www.audacityteam.org/ [Accessed 23 Aug. 2016].

Azenkot, S and Zhai, S 2012, Touch behavior with different postures on soft smartphone keyboards,
Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services - MobileHCI '12.

Brewster, S., Chohan, F. and Brown, L. (2007). Tactile feedback for mobile interactions. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07*, [online] pp.159-162. Available at: http://dl.acm.org/citation.cfm?id=1240649 [Accessed 23 Aug. 2016].

Butts, Cockburn. 2002. An evaluation of mobile phone text input methods. *In Proceedings of the Third Australasian conference on User interfaces* - Volume 7 (AUIC '02), Vol. 7. Australian Computer Society, Inc., Darlinghurst, Australia, Australia, 55-59.

Dagiene, V., Grigas, G. and Jevsikova, T. (2011). Comparison of Two 12-Key Assignments on Mobile Devices. *Electronics and Electrical Engineering*, 108(2).

Dunlop, M. and Levine, J. (2012). Multidimensional pareto optimization of touchscreen keyboards for speed, familiarity and improved spell checking. *Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems - CHI '12*, pp.2669-2678.

Dunlop, M. and Crossan, A. (2000). Predictive text entry methods for mobile phones. *Personal Technologies*, 4(2-3), pp.134-143.

Dunlop, M. and Taylor, F. (2009). Tactile feedback for predictive text entry. *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*, [online] pp.2257-2260. Available at: http://dl.acm.org/citation.cfm?id=1519046 [Accessed 23 Aug. 2016].

Freesound.org. (2016). *Freesound.org - Freesound.org*. [online] Available at: http://freesound.org [Accessed 23 Aug. 2016].

Garay-Vitoria, N. and Abascal, J. (2005). Text prediction systems: a survey. *Universal Access in the Information Society*, 4(3), pp.188-203.

Gong, J. and Tarasewich, P., 2005. Testing predictive text entry methods with constrained keypad designs. *Proceedings of Human Computer Interfaces International* (HCII 05).

Hoggan, E., Brewster, S. and Johnston, J. (2008). Investigating the effectiveness of tactile feedback for mobile touchscreens. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. [online] pp.1573-1582. Available at: http://dl.acm.org/citation.cfm?id=1357300 [Accessed 23 Aug. 2016].

James, C. and Reischel, K. (2001). Text input for mobile devices. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '01*, pp.365-371.

Kukich, K. (1992). Technique for automatically correcting words in text. *CSUR*, 24(4), pp.377-439.

Kwon, S., Lee, D. and Chung, M. (2009). Effect of key size and activation area on the performance of a regional error correction method in a touch-screen QWERTY keyboard. *International Journal of Industrial Ergonomics*, 39(5), pp.888-893.

(Louis P. SLOTHOUBER, Eric H. Davis, Michael K. Young, Jeffrey W. JOHNSTON, patent) Slothouber, L.P., Davis, E.H., Young, M.K. and Johnston, J.W., Biap, Inc., 2008. Context-dependent prediction and learning with a universal re-entrant predictive text input software component. U.S. Patent Application 12/116,917.

MacKenzie, I. and Isokoski, P. (2008). Fitts' throughput and the speed-accuracy tradeoff. *Proceeding of the twenty-sixth annual CHI conference on Human factors in computing systems - CHI '08*.

MacKenzie, I. and Soukoreff, R. (2003). Phrase sets for evaluating text entry techniques. *CHI '03 extended abstracts on Human factors in computing systems - CHI '03*, pp.754-755.

Media Playback | Android Developers. Developer.android.com. (2016). [online] Available at: https://developer.android.com/guide/topics/media/mediaplayer.html [Accessed 23 Aug. 2016].

NASA TLX Paper/Pencil Version. Humansystems.arc.nasa.gov. (2016). [online] Available at: http://humansystems.arc.nasa.gov/groups/tlx/paperpencil.html [Accessed 23 Aug. 2016].

Oulasvirta, A, Reichel, A, Li, W, Zhang, Y, Bachynskyi, M, Vertanen, K and Kristensson, P 2013, Improving two-thumb text entry on touchscreen devices, *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*.

Pavlovych, A. and Stuerzlinger, W., 2003, June. Less-Tap: A Fast and Easy-to-learn Text Input Technique for Phones*. In Graphics Interface* (Vol. 2003, pp. 97-104).

Rudchenko, D, Paek, T and Badger, E 2011, Text Text Revolution: A Game That Improves Text Entry on Mobile Touchscreen Keyboards, *Lecture Notes in Computer Science*, pp. 206-213.

Silfverberg, M., MacKenzie, I. and Korhonen, P. (2000). Predicting text entry speed on mobile phones. *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '00*, pp.9-16.

Tinwala, H. and MacKenzie, I. (2010). Eyes-free text entry with error correction on touchscreen mobile devices. *Proceedings of the 6th Nordic Conference on Human-Computer Interaction Extending Boundaries - NordiCHI '10*.

Yatani, K and Truong, K 2009, SemFeel, *Proceedings of the 22nd annual ACM symposium on User interface software and technology - UIST '09.*

Appendices

Appendix 1 – Consent form

The application will open with a consent form showing the following text along with a "Agree and continue" button:

Thank you for taking part in this study. The study forms part of research at The University of Strathclyde where we are investigating the effect of different feedback on text entry performance.

In the study you will be asked to enter around 100 short phrases on this mobile phone and complete a couple of short forms. The study will take around 1 hour. Performance and typing behaviour data is recorded but no personal data is recorded by the app. We hope to publish the results of the study in research papers and presentations – summary data and examples of typing may be included in these but no personal data will revealed.

You have the right to stop the study at any point by simply letting the researcher know you wish to stop.

Appendix 2 – Introductory Questionnaire

Investigation into directional tactile feedback for mobile text entry

Introductory Questionnaire

Date:_____________________

1.Age: ____

2.Gender:____

3. How many text messages a day on average do you send (using smartphone qwerty keyboards, drawing words eg swype) (circle your choice)

a) 20 or more

b) 10 or more

c) Between 5-10

d) Between 1-4

e) 2-3 a week

f) Fewer than 2-3 a week

4.) How long approximately have you been using text messaging on smartphones (utilising the above mentioned methods)?

____

Appendix 3 – NASA TLX

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Appendix 4 – End Questionnaire

Investigation into directional tactile feedback for mobile text entry

Survey

1. Were you able to notice a difference between the two vibrational feedback methods?
   - ☐ Yes
   - ☐ No

2. If you felt a difference between the two methods which did you prefer?
   - ☐ Stereo
   - ☐ Mono

3. Why?

4. Did you feel that the stereo vibrational feedback had an impact on your typing accuracy and speed?
   - ☐ Yes
   - ☐ No

5. If so, was this impact beneficial? Why/why not?

1.

6. How easy was the stereo feedback method to get used to?

1	2	3	4	5	6	7

Appendix 5 – Questionnaires and TLX forms filled out by users

MP1 first

Investigation into directional tactile feedback for mobile text entry

Introductory Questionnaire

Date: 10|8|16

1.Age: 37

2.Gender: Female

3. How many text messages a day on average do you send (using smartphone qwerty keyboards, drawing words eg swype) (circle your choice)

a) 20 or more

b) 10 or more

c) Between 5-10

d) Between 1-4

e) 2-3 a week

f) Fewer than 2-3 a week

4.) How long approximately have you been using text messaging on smartphones (utilising the above mentioned methods)?

5+ years.

1st MP1

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task: Text Entry	Date: 10/8/16

Mental Demand — How mentally demanding was the task? Very Low – Very High

Physical Demand — How physically demanding was the task? Very Low – Very High

Temporal Demand — How hurried or rushed was the pace of the task? Very Low – Very High

Performance — How successful were you in accomplishing what you were asked to do? Perfect – Failure

Effort — How hard did you have to work to accomplish your level of performance? Very Low – Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou? Very Low – Very High

2nd SP2

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task: Text Entry	Date: 10/8/16

Mental Demand — How mentally demanding was the task? Very Low – Very High

Physical Demand — How physically demanding was the task? Very Low – Very High

Temporal Demand — How hurried or rushed was the pace of the task? Very Low – Very High

Performance — How successful were you in accomplishing what you were asked to do? Perfect – Failure

Effort — How hard did you have to work to accomplish your level of performance? Very Low – Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou? Very Low – Very High

Investigation into directional tactile feedback for mobile text entry

Survey

1. Were you able to notice a difference between the two vibrational feedback methods?
   - ☒ Yes
   - ☐ No
2. If you felt a difference between the two methods which did you prefer?
   - ☒ Stereo
   - ☐ Mono
3. Why?

   With stereo I didn't much feel the difference in vibration but in sound so if it sounded different or "off" I could check for a mistake.
4. Did you feel that the stereo vibrational feedback had an impact on your typing accuracy and speed?
   - ☒ Yes
   - ☐ No
5. If so, was this impact beneficial? Why/why not?

   I think stereo helped with speed because you could type faster and listen for the sound to highlight a potential mistake. The only issue would be if you'd completely hit a wrong letter then you'd still need to visually check for mistakes but overall it felt faster.
6. How easy was the stereo feedback method to get used to?

1	2	3	4	5	6	7
					X	

EASY!

Sp2 first

Investigation into directional tactile feedback for mobile text entry

Introductory Questionnaire

Date: 10.08.16

1.Age: 19

2.Gender: Male

3. How many text messages a day on average do you send (using smartphone qwerty keyboards, drawing words eg swype) (circle your choice)

a) 20 or more

b) 10 or more

c) Between 5-10

d) Between 1-4

e) 2-3 a week

f) Fewer than 2-3 a week

4.) How long approximately have you been using text messaging on smartphones (utilising the above mentioned methods)?

2 years

Sp2 first

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task Mobile phone text input	Date 10.08.16

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Sp2 Second

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task Mobile phone text input	Date 10.08.16

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Investigation into directional tactile feedback for mobile text entry

Survey

1. Were you able to notice a difference between the two vibrational feedback methods?
   - ☑ Yes
   - ☐ No
2. If you felt a difference between the two methods which did you prefer?
   - ☐ Stereo
   - ☑ Mono
3. Why?

   It felt more comfortable when typing

4. Did you feel that the stereo vibrational feedback had an impact on your typing accuracy and speed?
   - ☑ Yes
   - ☐ No
5. If so, was this impact beneficial? Why/why not?

   It helped me identify when I had missed a key

6. How easy was the stereo feedback method to get used to?

1	2	3	4	5	6	7
						X

MP2

Investigation into directional tactile feedback for mobile text entry

Introductory Questionnaire

Date: 10/8/2016

1.Age: 22

2.Gender: Male

3. How many text messages a day on average do you send (using smartphone qwerty keyboards, drawing words eg swype) (circle your choice)

a) 20 or more

b) 10 or more

c) Between 5-10

d) Between 1-4

e) 2-3 a week

f) Fewer than 2-3 a week

4.) How long approximately have you been using text messaging on smartphones (utilising the above mentioned methods)?

6 years

1st

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
	Text entry	10/8/2016

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
	Text entry	10/8/2016

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Investigation into directional tactile feedback for mobile text entry

Survey

1. Were you able to notice a difference between the two vibrational feedback methods?
   - ☑ Yes
   - ☐ No

2. If you felt a difference between the two methods which did you prefer?
   - ☑ Stereo
   - ☐ Mono

3. Why?

Instant feedback on typing accuracy without looking at text input

4. Did you feel that the stereo vibrational feedback had an impact on your typing accuracy and speed?
   - ☑ Yes
   - ☐ No

5. If so, was this impact beneficial? Why/why not?

Yes, same as above. Tactile feedback is much easier to ~~[illegible]~~ notice and respond to rather than looking away from keyboard.

6. How easy was the stereo feedback method to get used to?

1	2	3	4	5	6	7
				✓		

SP1

Investigation into directional tactile feedback for mobile text entry

Introductory Questionnaire

Date: 10/08/16

1.Age: 23

2.Gender: Female

3. How many text messages a day on average do you send (using smartphone qwerty keyboards, drawing words eg swype) (circle your choice)

a) 20 or more

b) 10 or more

c) Between 5-10

d) Between 1-4

e) 2-3 a week

f) Fewer than 2-3 a week

4.) How long approximately have you been using text messaging on smartphones (utilising the above mentioned methods)?

3 years

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task App (no. 2)	Date 10/08/16

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Investigation into directional tactile feedback for mobile text entry

Survey

1. Were you able to notice a difference between the two vibrational feedback methods?
   - ☑ Yes
   - ☐ No
2. If you felt a difference between the two methods which did you prefer?
   - ☑ Stereo
   - ☐ Mono
3. Why?

The different notes gave me a better idea of how I was hitting the keys

4. Did you feel that the stereo vibrational feedback had an impact on your typing accuracy and speed?
   - ☑ Yes
   - ☐ No
5. If so, was this impact beneficial? Why/why not?

I feel like I typed a bit faster. I was frustrated with the mono task because a glitch kept deleting some letters which slowed me down, but that did not have anything to do with the sound or vibration.

6. How easy was the stereo feedback method to get used to?

1	2	3	4	5	6	7
	X					

NB.

I felt that the sound stood out to me more than the vibrations. The task demanded concentration so I was less aware of extra physical sensations, as I was concentrating on typing with one finger.

MP1

Investigation into directional tactile feedback for mobile text entry

Introductory Questionnaire

Date: 12/08/16

1.Age: 29

2.Gender: Female

3. How many text messages a day on average do you send (using smartphone qwerty keyboards, drawing words eg swype) (circle your choice)

a) 20 or more

b) 10 or more

c) Between 5-10

d) Between 1-4

e) 2-3 a week

f) Fewer than 2-3 a week

4.) How long approximately have you been using text messaging on smartphones (utilising the above mentioned methods)?

5 + years

first

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
	Text entry	12/08/16

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
	Text entry 2	12/08/16

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Investigation into directional tactile feedback for mobile text entry

Survey

1. Were you able to notice a difference between the two vibrational feedback methods?
   - ☑ Yes
   - ☐ No
2. If you felt a difference between the two methods which did you prefer?
   - ☐ Stereo
   - ☑ Mono
3. Why?

   The difference in stereo was distracting me more

4. Did you feel that the stereo vibrational feedback had an impact on your typing accuracy and speed?
   - ☑ Yes
   - ☐ No
5. If so, was this impact beneficial? Why/why not?

   I was slower cause I'd stop for a second but I didn't feel that it improved my accuracy

6. How easy was the stereo feedback method to get used to?

1	2	3	4	5	6	7
		✓				

hard — easy

MP2

Investigation into directional tactile feedback for mobile text entry

Introductory Questionnaire

Date: 8/12/2016

1.Age: 25

2.Gender: Male

3. How many text messages a day on average do you send (using smartphone qwerty keyboards, drawing words eg swype) (circle your choice)

a) 20 or more

b) 10 or more

c) Between 5-10

d) Between 1-4

e) 2-3 a week

f) Fewer than 2-3 a week

4.) How long approximately have you been using text messaging on smartphones (utilising the above mentioned methods)?

6 years

first

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task: Mono Text	Date: 8/12/16

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task: Stereo Text	Date: 8/12/16

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Investigation into directional tactile feedback for mobile text entry

Survey

1. Were you able to notice a difference between the two vibrational feedback methods?
   - ☑ Yes
   - ☐ No
2. If you felt a difference between the two methods which did you prefer?
   - ☑ Stereo
   - ☐ Mono
3. Why?

Stereo felt a little more realistic.

4. Did you feel that the stereo vibrational feedback had an impact on your typing accuracy and speed?
   - ☐ Yes
   - ☑ No
5. If so, was this impact beneficial? Why/why not?

I could mostly feel that the left sensor was activated more often, but it didn't really provide any improvement in how I typed.

6. How easy was the stereo feedback method to get used to?

1	2	3	4	5	6	7
				✓		

SP1

Investigation into directional tactile feedback for mobile text entry

Introductory Questionnaire

Date: 14.08.16

1.Age: 33

2.Gender: Male

3. How many text messages a day on average do you send (using smartphone qwerty keyboards, drawing words eg swype) (circle your choice)

a) 20 or more

b) 10 or more

c) Between 5-10

d) Between 1-4

e) 2-3 a week

f) Fewer than 2-3 a week

4.) How long approximately have you been using text messaging on smartphones (utilising the above mentioned methods)?

5 years

first

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
	text entry 1	14.08.16

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
	text test 2	14.08.16

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Investigation into directional tactile feedback for mobile text entry

Survey

1. Were you able to notice a difference between the two vibrational feedback methods?
   - ☒ Yes
   - ☐ No
2. If you felt a difference between the two methods which did you prefer?
   - ☒ Stereo
   - ☐ Mono

3. Why?

it broke up the monotony/habitual/reflex texting in which i tend to zone out and text away and send when done. by noticing different side vibrations i found myself briefly pausing and looking up at the line i was typing, and noticing my mistakes more than when texting without distraction/interruption.

4. Did you feel that the stereo vibrational feedback had an impact on your typing accuracy and speed?
   - ☒ Yes
   - ☐ No

5. If so, was this impact beneficial? Why/why not?

I found it was beneficial by giving a chance to focus more by subtle interruption to what is otherwise a very routine ~~task~~. mindless task.

6. How easy was the stereo feedback method to get used to?

1	2	3	4	5	6	7

MP2

Investigation into directional tactile feedback for mobile text entry

Introductory Questionnaire

Date: 16/08/2016

1.Age: 24

2.Gender: MALE

3. How many text messages a day on average do you send (using smartphone qwerty keyboards, drawing words eg swype) (circle your choice)

a) 20 or more

b) 10 or more

c) Between 5-10

d) Between 1-4

e) 2-3 a week

f) Fewer than 2-3 a week

4.) How long approximately have you been using text messaging on smartphones (utilising the above mentioned methods)?

6 years

First

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
	TEXT ENTRY	16/08/2016

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
	TEXT ENTRY	16/08/2016

Mental Demand — How mentally demanding was the task?

Very Low — Very High

Physical Demand — How physically demanding was the task?

Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?

Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?

Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?

Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?

Very Low — Very High

Investigation into directional tactile feedback for mobile text entry

Survey

1. Were you able to notice a difference between the two vibrational feedback methods?
   - ☐ Yes
   - ☒ No
2. If you felt a difference between the two methods which did you prefer?
   - ☐ Stereo
   - ☐ Mono
3. Why?

4. Did you feel that the stereo vibrational feedback had an impact on your typing accuracy and speed?
   - ☒ Yes
   - ☐ No
5. If so, was this impact beneficial? Why/why not?

No, the delayed feedback was annoying and I felt it slowed me down.
Although, It was useful to know if I managed to hit the button or not.

6. How easy was the stereo feedback method to get used to?

1	2	3	4	5	6	7
				X		

SPe

Investigation into directional tactile feedback for mobile text entry

Introductory Questionnaire

Date: 17/08/16

1.Age: 26

2.Gender: Male

3. How many text messages a day on average do you send (using smartphone qwerty keyboards, drawing words eg swype) (circle your choice)

a) 20 or more
b) 10 or more
c) Between 5-10
d) Between 1-4
e) 2-3 a week
f) Fewer than 2-3 a week

4.) How long approximately have you been using text messaging on smartphones (utilising the above mentioned methods)?

7 years

First

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
	Text Entry	17/08/16

Mental Demand — How mentally demanding was the task?
Very Low — Very High

Physical Demand — How physically demanding was the task?
Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?
Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?
Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?
Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?
Very Low — Very High

Figure 8.6

NASA Task Load Index

Hart and Staveland's NASA Task Load Index (TLX) method assesses work load on five 7-point scales. Increments of high, medium and low estimates for each point result in 21 gradations on the scales.

Name	Task	Date
	Text Entry	17/08/16

Mental Demand — How mentally demanding was the task?
Very Low — Very High

Physical Demand — How physically demanding was the task?
Very Low — Very High

Temporal Demand — How hurried or rushed was the pace of the task?
Very Low — Very High

Performance — How successful were you in accomplishing what you were asked to do?
Perfect — Failure

Effort — How hard did you have to work to accomplish your level of performance?
Very Low — Very High

Frustration — How insecure, discouraged, irritated, stressed, and annoyed wereyou?
Very Low — Very High

Investigation into directional tactile feedback for mobile text entry

Survey

1. Were you able to notice a difference between the two vibrational feedback methods?
   - ☒ Yes
   - ☐ No
2. If you felt a difference between the two methods which did you prefer?
   - ☐ Stereo
   - ☒ Mono
3. Why?

   I found that just having a tone when touching a key helped assure me that I had hit the key. So I double checked my entry less often.

4. Did you feel that the stereo vibrational feedback had an impact on your typing accuracy and speed?
   - ☒ Yes
   - ☐ No
5. If so, was this impact beneficial? Why/why not?

   It made me realise I was leaning more to one side than the other and adjust my hand positioning accordingly.

6. How easy was the stereo feedback method to get used to?

1	2	3	4	5	6	7
				●		