Usability of Visibly Adaptive Smartphone keyboard layouts

Apostolos Gkoumas
University of Patras
Rio, 26500
Greece
gkoumas@ceid.upatras.gr

Andreas Komninos
University of Strathclyde
26 Richmond St., Glasgow
G1 1XH, UK
andreas@komninos.info

John Garofalakis
University of Patras / RACTI
Rio, 26500
Greece
garofala@ceid.upatras.gr

ABSTRACT

We present two touchscreen smartphone keyboard interface designs which dynamically alter key sizes as the user types and highlight likely key targets, based on word completion predictions. Although we find a speed reduction and no accuracy advantage with these methods over standard QWERTY, the text entry experience on mobile devices with these designs is significantly better for users.

CCS Concepts

• **Human-centered computing ~ Keyboards** • **Human-centered computing ~ Text input** •**Human-centered computing ~ Empirical studies in ubiquitous and mobile computing**

Keywords

Mobile text entry; touchscreen keyboards

1. INTRODUCTION

Mobile text entry is a critical component of the overall user experience in taking advantage of the wide variety of services that modern smartphones offer. The process of inputting text into smartphones is compounded by a range of factors. Small screens make for small key targets that are difficult to hit, particularly when keyboard layouts such as QWERTY are used for input. Ergonomic considerations such as the user's thumb or finger width further aggravate the problem of small targets. Furthermore, the partial occlusion of the screen area where the keyboard is drawn by the user's fingers can introduce uncertainty on where the next target is and may limit the visibility of touch feedback that lets the user know whether or not they have hit the right target. In the past, several approaches to addressing some of these issues have been covered in literature, using various methods such as automatically resolving users' intentions regarding key presses through word and touch modelling. Other approaches have focused on alternative layouts or dynamic adaptations of familiar layouts such as QWERTY. In our paper, we focus on two concepts: providing visual feedback to users in order to assist them in detecting likely target keys and dynamically adapting the keyboard layout to offer better support for desired input.

2. PREVIOUS RELATED WORK

Given Fitts' Law, the size of touchscreen keyboards is a critical factor in the ability of users to input text quickly and accurately. In [14] touchscreen text entry on a relatively large surface was examined. It was found that input speed is proportional to the keyboard's dimension, and error rate is inversely proportional to these. These results were later affirmed by [8], who also highlight that smaller targets also result in users requiring longer to find the next desirable target. These studies were based on the use of QWERTY layouts of different key dimensions. However, a further study on various keyboard layouts on mobile screens [10] showed that user familiarity with the keyboard layout also has a considerable effect on input speed, since users were able to perform faster with the QWERTY layout than other optimized keyboards. The effect of familiarity is also shown in [9] where users were able to outperform QWERTY using an optimized keyboard layout but only after prolonged exposure to the keyboard. Users' physical characteristics such as thumb size have also been shown to affect input performance when considered in relation to keyboard keys [2] (e.g. users with "fat" fingers exhibit worse performance in keyboards with smaller keys).

2.1 Dynamic adaptation of key sizes

To address some of the problems identified above, other literature demonstrates approaches in dynamically adapting keyboard layouts and take advantage of the entire screen space. In [1], the authors experimented with a virtual keyboard that dynamically altered the size of up to four keys based on predictions on the users' intended words. The keys are scaled linearly between 11.11% and 44.44% of a standard key size (18x16px) based on the raked probability of them being the next ones to be entered (considering previous input). The authors performed a study with a limited sample size (9 participants) who entered four sentences on a PDA device using a stylus (and not their actual fingers) using a simple QWERTY and one-key and four-key size variants of the designed method. In their study, altering the size of four likely letters resulted in an improved input speed and reduced error rates, however the paper does not report any statistical significance tests for these findings. In another approach reported in [5], the familiar QWERTY keyboard is dynamically re-arranged in a Voronoi-like manner, based on the user's touch coordinate history. This method was found to show similar performance to the QWERTY keyboard. However, research by [7] hints that users may experience discomfort with continuous adaptation of the layout. As such, other approaches (e.g. [6, 3]) dynamically adapt key sizes based on word completion predictions, informed by users' typing history and corrections, but

PCI '16, November 10 - 12, 2016, Patras, Greece

ACM 978-1-4503-4789-1/16/11...$15.00
DOI: http://dx.doi.org/10.1145/3003733.3003743

do not alter the keyboard's visual layout, so as to not distract users.

2.2 Visual feedback without key size adaptation

To address the issue of difficulty in searching for key targets, in [12], the authors performed a comparative evaluation of alternative keyboard designs, including highlighting of next four likely keys (derived from a list of word completion predictions), resizing of likely keys, and enlarging the touch area for a likely key without altering its appearance. They performed comparative lab tests with young adults on a tablet device and found that overall the alternative designs did not manage to improve the typing speeds over simple QWERTY. However, all three novel methods resulted in fewer input errors than simple QWERTY. Young adults however declared strong preferences for the static visual representation of keys. The authors concluded that further research with designs that combine the aforementioned techniques is needed. In [13] the same researchers focused on the impact of key highlighting with older adults, again on tablet devices. It was found that this technique decreased input speed slightly and caused fewer insertion (but slightly more omission and substitution) errors compared to QWERTY. Participants however rated this method highly and though researchers attributed the slower input speed to the distraction possibly provided by the key highlighting, no participant mentioned this as a concern.

A study reported in [11] provides a possible insight to the combination of key resizing and simultaneous highlighting. The authors performed a comparative evaluation of five commercially available input methods including QWERTY, and ThickButtons[1]. The ThickButtons method implements the research recommendation of [12] and simultaneously enlarges likely keys (reducing the size of other keys), and also alters likely key colour. The reduction of unlikely key sizes depends on how many likely keys are on the same row, hence the more likely keys in a row, the smaller the rest of the keys in that row become. The authors carried out a lab evaluation using just six participants using a smartphone, each participant carrying out four input tasks with each keyboard type, and measured the total keypresses and time taken to complete tasks (thereby computing input speed). In all tasks, the ThickButtons method was found to be consistently the worst performing. No evaluation of input accuracy is reported.

2.3 Weaknesses in related work

The aforementioned studies show some hints that visual adaptation of keyboard layouts may offer hinderances to participants. The studies on visible dynamic key resizing and key highlighting, suffer from a very limited number of participants and also a very limited number of input tasks [1, 11], and are hence methodologically weak. The best insights seem to come from [12] and [13], although they only tested using tablet devices and not smartphones. Additionally, none of these two studies investigates the combination of key resizing and key highlighting. Hence our paper aims at examining visual adaptation of touchscreen keyboards, on smartphone-size devices and with a younger adult population. In this paper thus, we explore the combination of key resizing, with key highlighting, using an adequate sample size and performing our evaluation on a smartphone.

3. THE ADAPTIVE KEYBOARD

To fulfil our goal, we built an adaptive keyboard that works in two modes (Figure 1): Uniform key resizing (UKR) and Neighbouring key resizing (NKR), as explained below. To determine candidate keys for visual adaptation, we used the open-source OpenAdapTxt[2] predictive engine to obtain word completion suggestions as the user typed. From the engine's suggestions list, we considered the first three words.

3.1 Key size adaptation

Uniform key resizing (UKR) is the visible key resizing method used in past literature, where all the unlikely keys are evenly resized depending on the amount of space left in a row after the likely keys have been resized (e.g. [11, 12, 13]). In addition to this, we developed an alternative resizing method (NKR), in which we selected an uneven resizing of the unlikely keys by detracting only from the size of keys neighbouring a likely key, while leaving the rest of the keys in their standard size. We extended the likely key's width by 20%, while simultaneously detracting 5% from the width of its left immediate neighbouring key and 15% from the width of its right neighbor. This design aimed at right-handed users where the occlusion of keys by the user's thumb occurs on the right, hence making the user unsure where their touch is going to "hit". In [13] it is also recommended that touches are "shifted" to the top and left (for right-handed users), precisely because of this phenomenon. A further advantage of this approach was that it limits the resizing of keys to just those neighbouring a candidate. Hence, when the suggestion engine provides a candidate other than the one that the user intends to hit, the resizing effects are limited to just those around the wrong suggestion and not all keys in the same row. Finally we should note that we also extended the likely key's height in both UKR and NKR by 20%, effectively "eating into" the blank space separating the keys but without causing key overlap.

3.2 Cumulative Key Highlighting

Since we aim to add to the literature by exploring the combination of visible key resizing with key highlighting, for the highlighting of candidate keys, we added a semi-transparent white layer to each key. If a key was a candidate in the same word position in more than one suggestion (e.g. the user types "t" and the suggestions were "test, testing, toast"), we place two semi-transparent overlays on the key "e" and one on "o", highlighting the "e" more prominently than the "o". Candidates are calculated even before the first key of a word has been pressed (in which case they are typical sentence starting words, e.g. "the", "a" etc.).

---

[1] ThickButtons Application: https://play.google.com/store/apps/details?id=com.thickbuttons

[2] OpenAdapTxt library http://openadaptxt.sourceforge.net/

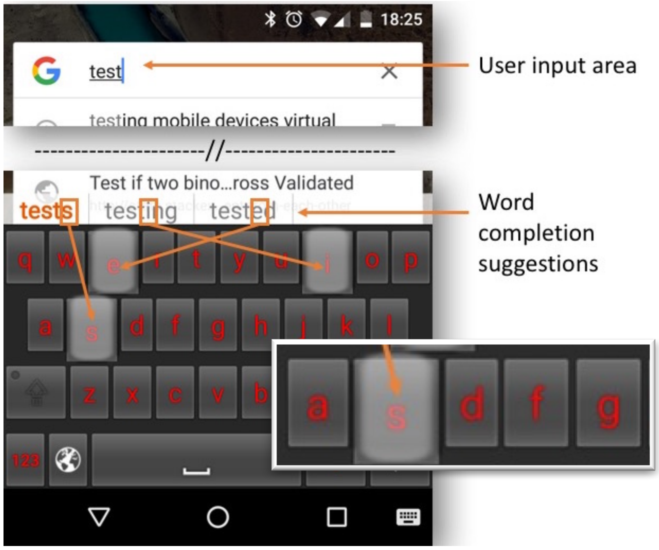

Figure 1. The NKR keyboard design. As is visible, only the keys neighbouring a likely key are altered in size.

4. EXPERIMENT AND RESULTS

4.1 Experiment setup

To test our keyboard designs, we performed a laboratory test with 20 Computer Science students (aged 20-25, 8 female), all right-handed. All participants self-reported as being experienced smartphone users and using the QWERTY keyboard daily on their smartphones. We asked our participants to copy 14 sentences from the Enron Memorable Phrase Set using a plain QWERTY keyboard, UKR and NKR, on the same Nexus 4 mobile device running Android 5.1 (within-subject experiment design). The first three phrases were quite short (22 characters each), the next 5 were longer (37.4 characters long, SD=0.55) and the final six were approximately 58.2 characters long (SD=3.73).

Each participant performed input with a different keyboard sequence (counterbalancing), in order to avoid any learning effects. We recorded, for each phrase, the calculated WPM (input speed) and the number of backspaces pressed by the user. After each keyboard session, we asked participants to fill in a UEQ questionnaire, in order to measure their experience during use. We also asked participants to fill in a NASA-TLX questionnaire. Finally, at the end of each session, we asked participants to respond to four personal evaluation questions, with responses recorded on the Likert scale (1=completely disagree, 5=completely agree).

4.2 Choice of statistical tests

All reported statistical tests below (quantitative and qualitative) were chosen based on the distributions of variables (calculated using a Shapiro-Wilk test), using either Wilcoxon signed rank or paired sample T-tests as appropriate to perform pairwise comparisons.

4.3 Quantitative measures results

4.3.1 Input speed

Participants were fastest with the simple QWERTY keyboard (m=24.2WPM, stdev=6.17WPM). Their input speed was 14% slower with UKR (m= 20.73WPM, stdev=5.21WPM) and 16% slower with NKR(m= 20.20WPM, stdev=5.12WPM).

We found that the observed differences were statistically significant comparing QWERTY to UKR (Z=-3.211, p<0.01, Wilcoxon) and QWERTY to NKR (Z=-3.472, p<0.01, Wilcoxon). We did not find a statistically significant difference between UKR and NKR (Z=-0.933, p=0.351).

These results show that even though our designs were slower, participants were still able to maintain relatively good input speeds with both. The lack of statistically significant differences between UKR and NKR also shows that the performance penalty in input speed can be attributed to the need for assessing changes in the keyboard display, an effect which is supported by previous literature.

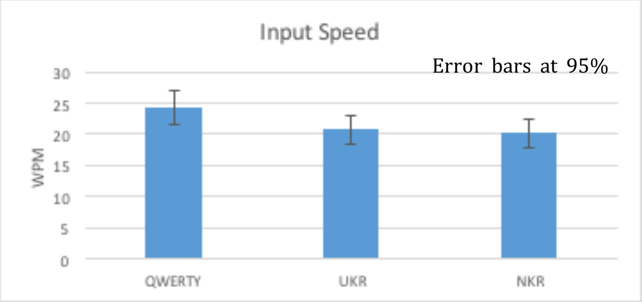

Figure 2. Input speed with the three keyboards

4.3.2 Input accuracy and errors

Although we expected that a larger target would improve our participants' accuracy in tapping keys, our results indicate that users were equally accurate in tapping the desired keys during text composition with all keyboards participants, since they exhibited a similar amount of backspace use with all keyboards (Figure 3). On average, participants used the backspace key 4.52 times per task (stdev=2.16) using QWERTY, 4.22 times per task (stdev=2.94) with UKR and 4.14 times per task (sd-3.48) with NKR. We didn't find any statistically significant differences when comparing QWERTY to UKR (Z=-1.139, p=0.255, Wilcoxon) or QWERTY to NKR (Z=-0.971, p=0.332, Wilcoxon). There was also no statistically significant difference between UKR and NKR (t(19)=0.189, p=0.852, t-test). Of course, backspace use is affected by the participants' ability to notice a mistaken key pressed and thus to correct it.

To calculate the accuracy of submitted text (i.e. errors that were not detected and thus not corrected by backspacing), we computed the Levenshtein string difference between the requested text and the actual submitted text, removing all capitalization except that at the start of the submitted sentence, and any punctuation at the end of the submitted text, as our requested text did not include any (Figure 3).

As regards the accuracy of submitted text, on average, the Levenshtein distance in each task was 2.07 (stdev=0.25) for QWERTY, 2.27 (stdev=0.79) for UKR and 1.98 (stdev=0.07) for NKR. The difference in accuracy of the submitted input was not statistically significantly when comparing QWERTY to UKR (Z=-0.751, p=0.453, Wilcoxon) and QWERTY to NKR (Z=-1.926, p=0.054, Wilcoxon). However, we did find a statistically significant difference between UKR and NKR (Z=-2.587, p<0.01, Wilcoxon).

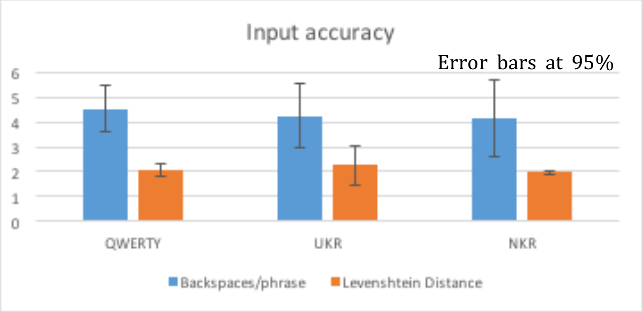

Figure 3. Input accuracy metrics with the three keyboards

These results show that limiting the shrinking of the keys to just the ones neighbouring a likely key, resulted in better final input accuracy than with the even resizing of all keys. This effectively limits the negative effect of the shrinking of all other targets, when the suggestion engine fails to produce good key candidates, resulting in performance that is on par with the simple QWERTY keyboard.

4.4 Qualitative evaluation results

4.4.1 Subjective evaluation – UEQ

The UEQ measures user experience in six axes – Attractiveness, Perspicuity, Efficiency, Dependability, Stimulation and Novelty (Figure 4).

In terms of attractiveness, our new designs outperformed the traditional QWERTY keyboard (QWERTY-UKR Z=-3.509, p<0.01, QWERT-NKR Z=-3.698, p<0.01, Wilcoxon). Participants also thought that the NKR design was more attractive than UKR (UKR-NKR Z=-3.705, p<0.01, Wilcoxon). For perspicuity, there was no statistically significant difference between the three keyboards (QWERTY-UKR Z= -0.650, p=0.516, QWERT-NKR Z=-1.946, p=0.052, UKR-NKR Z=-1.495, p=0.135, Wilcoxon).

Our new designs were deemed more efficient than the QWERTY keyboard (QWERTY-UKR Z=-3.286, p<0.01, QWERT-NKR Z=-3.565, p<0.01, Wilcoxon). Participants also found the NKR design was more efficient than UKR (UKR-NKR Z=-2.687, p<0.01, Wilcoxon). Both our new designs again were found to be more dependable than the QWERTY keyboard (QWERTY-UKR Z=-2.982, p<0.01, QWERT-NKR Z=-3.199, p<0.01, Wilcoxon). Participants also found the NKR design was more dependable than UKR (UKR-NKR Z=-2.608, p<0.01, Wilcoxon).

On the axis of stimulation, our new designs outperformed the QWERTY keyboard (QWERTY-UKR Z=-3.260, p<0.01, QWERT-NKR Z=-3.727, p<0.01, Wilcoxon). Participants also found the NKR design was more stimulating than UKR (UKR-NKR Z=-3.539, p<0.01, Wilcoxon). Finally, with regard to novelty, our new designs fared better than the QWERTY keyboard (QWERTY-UKR Z=-3.683, p<0.01, QWERT-NKR Z=-3.924, p<0.01, Wilcoxon). Participants also found the NKR design was more novel than UKR (UKR-NKR Z=-2.987, p<0.01, Wilcoxon).

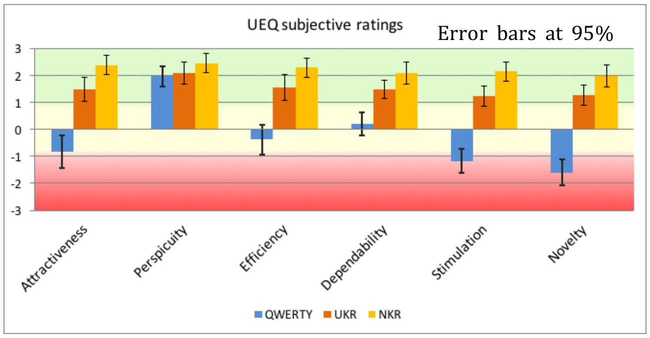

Figure 4. UEQ subjective evaluations

In summary, the user experience with both our new designs was better on all fronts except perspicuity. Participants also seemed to rate their experience with the limiting key size reductions to candidate neighbours higher than all others.

4.4.2 Subjective evaluation – NASA TLX

As mentioned, after each keyboard session, we asked participants to fill in a NASA-TLX questionnaire (Figure 5).

On the scale of mental demand, we did not find any statistically significant differences (QWERTY-UKR t(19)=-1.371, p=0.186, QWERT-NKR t(19)=1.143, p=0.267, UKR-NKR t(19)=1.630, p=0.119, paired sample t-tests).

In terms of perceived physical demand, again we did not find any statistically significant differences (QWERTY-UKR t(19)=0.767, p=0.453, QWERT-NKR t(19)=1.759, p=0.095, UKR-NKR t(19)=1.718, p=0.102, paired sample t-tests).

On the scale of temporal demand, a statistically significant difference was found between QWERTY and NKR (t(19)=2.334, p<0.05, paired sample t-tests) but not other comparisons (QWERTY-UKR t(19)=1.690, p=0.107, UKR-NKR t(19)=1.697, p=0.106, paired sample t-tests).

For perceived performance, a statistically significant difference was found between QWERTY and NKR (t(19)=3.347, p<0.01, paired sample t-tests) but not in other comparisons (QWERTY-UKR t(19)=1.372, p=0.186, UKR-NKR t(19)=2.009, p=0.059, paired sample t-tests).

On the effort scale, a statistically significant difference was found between QWERTY and UKR (t(19)=2.199, p<0.05, paired sample t-tests) and QWERTY and NKR (t(19)=3.601, p<0.01, paired sample t-tests) but not between UKR and NKR (QWERTY-UKR t(19)=1.502, p=0.150, paired sample t-tests). Finally, in terms of reported frustration, we found statistically significant differences in all comparisons (QWERTY-UKR t(19)=3.293, p<0.01, QWERT-NKR t(19)=4.667, p<0.01, UKR-NKR t(19)=2.662, p<0.05, paired sample t-tests).

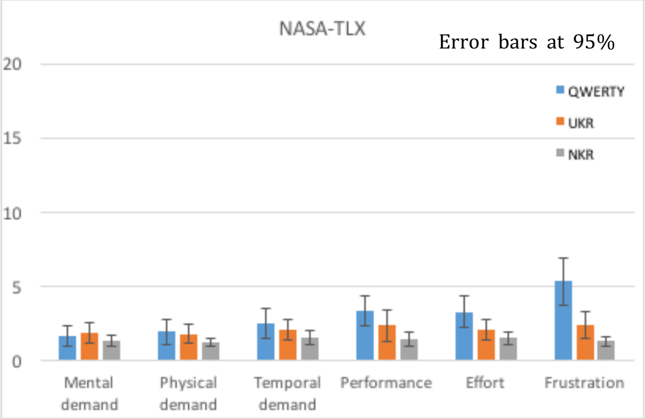

Figure 5. NASA-TLX subjective evaluations

In summary, participants didn't feel that the mental demand during input was any different with our new designs and in all cases, the tasks had a very low mental, temporal, physical and effort demands with all keyboards. Frustration was also generally low; however, our participants were least frustrated during input with the new designs (even more so with NKR). In terms of performance, again they felt that their performance was generally very good (a lower score on this scale is better) but, they felt they performed better with NKR.

In a sense, the findings on performance, effort and frustration seem to contrast our quantitative results, where we found that they were faster in typing and exhibited the same level of accuracy with the simple QWERTY keyboard. From these results it becomes apparent that our participants considered performance to be a function of input speed, accuracy and also the ease with which they performed input. Given that the input speed with the new designs was lower than QWERTY but still quite high, it appears that our new designs and particularly NKR made it easier for participants to carry out the input tasks by directing their attention to the likely characters and reducing the effort of searching and tapping. NKR also contributed to the participants' perceived performance positively compared to UKR, as it reduced the negative impact of a suggestion engine recommending the wrong keys.

4.4.3 Subjective evaluation – personal questions

At the end of each session, we asked participants to respond to the following statements for each keyboard:

Q1. It was easy for me to learn to use this keyboard ; Q2. I would use this keyboard on a daily basis

For UKR and NKR we also asked the following two questions to explore the perceived usefulness of the highlighting and key resizing features:

Q3. Highlighting the keys helped me to easily locate the keys I was looking for ; Q4. Changing the key sizes helped me to easily locate the keys I was looking for.

From the responses to the first two questions (Q1, Q2), we note that our participants found all keyboards easy to learn (strongly agree or agree QWERTY 95%, UKR 95%, NKR 95%), while being slightly more positive towards NKR (strongly agree 80% vs. 55% for the other two). Their views on whether they would use each keyboard on a daily basis were varied for QWERTY, in which 30% agreed or strongly agreed with this statement and 50% disagreed or strongly disagreed. For UKR 40% strongly agreed, 40% agreed, 10% were neutral and 10% disagreed. In contrast, 70% of participants strongly agreed that they would use NKR, 20% agreed and 5% were neutral or disagreed respectively.

More interestingly, the remaining two questions (Q3, Q4) related to how useful our participants thought the key resizing and key highlighting were in the UKR and NKR. In UKR 30% of participants strongly agreed that the key highlighting was helpful, 60% agreed and 10% disagreed. At the same time, 25% strongly agreed that the uniform resizing of the keys was helpful, 55% agreed and 20% were neutral.

In NKR, 75% of participants strongly agreed that highlighting was helpful, with 20% agreeing and 5% neutral. The resizing of neighbouring keys was strongly found to be helpful also by 80% of participants, while the remaining 20% also agreed that it was a helpful feature. Comparing these results, our participants exhibited a strong willingness to use NKR as a daily keyboard. In NKR, a strong majority found that both the highlighting and key resizing features were very helpful, compared to UKR where both were perceived as just helpful by the majority of users.

5. CONCLUSIONS

In this paper, we address a gap in literature on visibly dynamic keyboard layouts. Previous work suffered from either limited participant numbers and tasks, which made results inconclusive, or in the fact that it didn't address the performance of visibly dynamic keyboard layouts on smartphones. Additionally, the current state of the art in this area resulted in a recommendation that key resizing and key highlighting should be examined as a combined design. We presented thus our investigation in two designs combining visible key resizing and key highlighting, comparing these with a traditional QWERTY keyboard on a smartphone, using a balanced lab trial. Both designs are based on previous literature [9], though the second design (NKR) adds novelty in limiting key resizing to candidate and adjacent keys only.

Our results indicate that the dynamic resizing and highlighting of keys on soft keyboard on smartphones causes users to type more slowly than a simple QWERTY keyboard, maintaining however, an input speed that is satisfactory and above 20WPM. We also didn't find any evidence that the combination of resizing and highlighting keys offers any advantages in input accuracy over QWERTY, although there are caveats to this finding. Our participants were all heavy smartphone users and accustomed to the layout of QWERTY, thus had little to benefit from the indication of desired key position in the keyboard. Additionally, users seemed equally able to hit target keys, because the resizing of the likely and unlikely keys was modest. However, we noted that NKR offered better accuracy compared to UKR, a finding which can be attributed to limiting the impact of bad key suggestions by the engine to just the neighbouring keys.

Despite these findings, both the UKR and NKR techniques involving key resizing and highlight offer a better user experience than QWERTY, as reported by our participants, in terms of less effort, less frustration, better perceived efficiency and dependability. We believe that with training, users may be able to overcome any effects from the changing of the keyboard layout and achieve speeds on par with those of simple QWERTY keyboards. Since participants indicated a strong willingness to use these keyboards on a daily basis (especially NKR), it would be worth expanding our work in longitudinal trials and with other special populations such as older adults with cognitive or motor impairments, or users that are not very familiar with QWERTY. We would also like to try our design with miniscule screens such as smartwatches, where target size is a significant problem and compare outcomes with keystroke level modelling results.

6. REFERENCES

1. Khaldoun Al Faraj, Mustapha Mojahid and Nadine Vigouroux. 2009. BigKey: A Virtual Keyboard for Mobile Devices. *Lecture Notes in Computer Science*, 3–10. DOI:10.1007/978-3-642-02580-8_1
2. Vimala Balakrishnan and H. P. Paul. 2008. A study of the effect of thumb sizes on mobile phone texting satisfaction. *J. Usability Studies* 3, 3, 118-128.
3. Tyler Baldwin and Joyce Chai. 2012. Towards online adaptation and personalization of key-target resizing for mobile devices. In *Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*. ACM, New York, NY, USA, 11-20. DOI:10.1145/2166966.2166969
4. Fitts 1954. The information capacity of the human motor system in controlling the amplitude of movement. Journal of experimental psychology. 47, 6 (1954), 381
5. Kentaro Go and Yuki Endo. 2007. CATKey: customizable and adaptable touchscreen keyboard with bubble cursor-like visual feedback. In *Human-Computer Interaction 2007*. Springer Berlin Heidelberg, 493-496. DOI:10.1007/978-3-540-74796-3_47
6. Asela Gunawardana, Tim Paek, and Christopher Meek. 2010. Usability guided key-target resizing for soft keyboards. In *Proceedings of the 15th international conference on Intelligent user interfaces* (IUI '10). ACM, New York, NY, USA, 111-118. DOI:10.1145/1719970.1719986
7. Johan Himberg, Jonna Häkkilä, Petri Kangas, and Jani Mäntyjärvi. 2003. On-line personalization of a touch screen based keyboard. In *Proceedings of the 8th international conference on Intelligent user interfaces* (IUI '03). ACM, New York, NY, USA, 77-84. DOI:10.1145/604045.604061
8. Jeong Ho Kim, J. H., Lovenoor S. Aulck, Ornwipa Thamsuwan, Michael C. Bartha and Peter W. Johnson. 2014. The effect of key size of touch screen virtual keyboards on productivity, usability, and typing biomechanics. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 56(7), 1235-1248. DOI:10.1007/978-3-642-39182-8_28
9. I. Scott MacKenzie and Shawn X. Zhang. 1999. The design and evaluation of a high-performance soft keyboard. In Proceedings of the SIGCHI conference on Human Factors in Computing Systems (CHI '99). ACM, New York, NY, USA, 25-31. DOI:10.1145/302979.302983
10. I. Scott MacKenzie, Shawn X. Zhang, S. X. and William R. Soukoreff. 1999. Text entry using soft keyboards. *Behaviour & information technology*, 18 (4), 235-244, DOI:10.1080/014492999118995.
11. Tom Page. 2013. Usability of text input interfaces in smartphones. *Journal of Design Research*, 11(1), 39. DOI:10.1504/jdr.2013.054065
12. Élvio Rodrigues, Micael Carreira and Daniel Gonçalves. 2013. Improving Text-Entry Experience on Tablets. In *Proceedings of the 5th Portuguese Conference on Human-Machine Interaction*. Vila Real, Portugal.
13. Élvio Rodrigues, Micael Carreira and Daniel Gonçalves. 2014. Improving Text-Entry Experience for Older Adults on Tablets. In *Universal Access in Human-Computer Interaction. Aging and Assistive Environments*. Springer International Publishing, 167-178. DOI:10.1007/978-3-319-07446-7_17
14. Andrew Sears, Doreen Revis, Janet Swatski, Rob Crittenden and Ben Shneiderman. 1993. Investigating touchscreen typing: the effect of keyboard size on typing speed. *Behaviour & Information Technology*, *12*(1), 17-22. DOI: 10.1080/01449299308924362