Effects of Keyboard Background on Mobile Text Entry

Ashish Yadav, Ahmed Sabbir Arif
Human-Computer Interaction Group
University of California, Merced
Merced, CA 95343 USA
{ayadav6, asarif}@ucmerced.edu

ABSTRACT
This paper presents results of a comparative study that investigated the effects of different types of keyboard backgrounds (themes) on actual and perceived text entry performance, in terms of speed and accuracy. Two color and two image backgrounds were compared with the default Google Android keyboard. Results revealed that keyboard background does not affect actual performance, however has a significant effect on perceived performance. Most participants felt that image backgrounds, regardless of whether they were pre or self-selected, affected their speed and accuracy. This suggests that it may be possible to enhance one's text entry experience simply by designing an effective keyboard theme. This paper concludes with reflections on how these findings could benefit text entry researchers and keyboard developers.

Author Keywords
Text entry; virtual keyboard; Qwerty; smartphone; mobile; touchscreen; background; color; theme.

CCS Concepts
• **Human-centered computing→Text input;** *Empirical studies in interaction design*

INTRODUCTION
The ubiquity of mobile text entry has resulted in the availability of numerous third-party virtual keyboards for tablets and smartphones. While many of these keyboards are targeted at languages other than English and special user groups, such as the disabled, children, and elderly, or attempt to improve on the state-of-the-art by incorporating new features or predictive systems, many simply customize the look-and-feel of the Qwerty layout. These keyboards usually do not alter the factors that have been proven to affect text entry performance, such as keyboard and key sizes [11,12,20–23], instead use different backgrounds and colors to attract users. Some of these keyboards also enable users to select background images or colors of their choice (e.g., [26–30]). Recently, the default Android OS keyboard has also enabled this feature [31]. However, no data is available on whether customizing keyboard background impacts text entry performance and user preference. This paper attempts to answer this question through an empirical study.

MUM '18, November 25–28, 2018, Cairo, Egypt
© 2018 Association for Computing Machinery.
ACM ISBN 978-1-4503-6594-9/18/11…$15.00
https://doi.org/10.1145/3282894.3282905

RELATED WORK
Many have investigated the effects of color in user interfaces. They recommended using simple color schemes in user interfaces since its difficult for users to develop an effective mental model when they are overwhelmed or confused by too many colors competing for attention [16,17,25].

An early work reported that the degree of visual distinction between different tasks impacts the extent of possible interference and perceived difficulty [8]. It showed that using opposing contrast level outlines to encircle objects and fonts (e.g., black border outlines for white objects) improves visibility and distinctiveness of items in transparent interfaces. Some keyboard themes attempt capitalize on this. A different work found out that the search time for finding an item decreases *"if the color of the item is known ahead of time*, and *if the color only applies to that item"* [18,25].

Numerous works have stressed the importance of using proper color codes (picked based on common practices and cultural usage) in user interfaces to reduce misinterpretations and incorrect responses [16,25]. Some have also studied the effects of color on performance. One work reported that the color of computer screens can affect visual task performance [9], while another identified a relationship between personality and how users select and organize their desktop wallpapers [14]. A different work [2] suggests that translucency does not affect text entry speed or accuracy. Yet, to our knowledge, no prior work has examined the effects of different virtual keyboard backgrounds on input performance and preference.

Motivation
This research is motivated by the following considerations. First, when comparing a novel or improved virtual keyboard with the state-of-the-art, or studying human interactions with virtual keyboards, researchers spend a substantial amount of time and effort in maintaining a visual resemblance between the prototype and the baseline keyboard (e.g., [1,4,13,24]). This due to the consideration that a different look-and-feel can be a "confound" in the evaluation. This work will inform them whether this rigorous process is really necessary. Second, it will also inform practitioners about whether

enabling different themes for virtual keyboard is indeed a good idea.

AN EXPERIMENT

We conducted a user study to test the following hypotheses. ***H1***) Virtual keyboard background affects text entry performance, in terms of speed and accuracy. ***H2***) Virtual keyboard background influence perceived text entry performance, in terms of speed and accuracy.

Apparatus

We used an Apple iPhone 7, 138.3×67.1×7.1 mm, 138 g for the study (Figure 1). It ran on the Apple iOS 11.2.6 at 326 ppi. We used Gboard, the default Android keyboard [31], in the study since it enables users to apply different colors and images as background, which they could select from either the theme library or their personal image galleries. All predictive features of the keyboard were disabled, including word prediction and autocorrection, to eliminate a potential confounding factor (since some participants may heavily rely on the predictive features, while the others may not). Text entry performance was recorded through WebTEM [3], which is a freely available cross-platform Web application for recording text entry performance metrics.

Participants

Twelve volunteers (5 female, 7 male) participated in the user study. Their age ranged from 22 to 30 years, average 25.75 years (SD = 2.4). They all had normal visual acuity and none of them were color blind. All were experienced smartphone users and had an average of 9.6 years of experience in mobile text entry (SD = 2.5). However, none of them had any experience with the keyboard themes used in the study. They all received a small compensation for volunteering.

Design

The study used a within-subjects design, where the independent variable was the keyboard background and dependent variables were the performance metrics and user responses. There were 5 background conditions: default, warm *color*, *cool color*, *preselected image*, and *user-selected image*. We included two color conditions based on prior research that reported that color affects psychological functioning of humans [7,9]. All conditions enabled key borders. All participants started with the default condition, then the other conditions were counterbalanced. However, we anticipated a negligible effect of order since all our participants were experienced virtual Qwerty users. In each condition, participants entered fifteen random English phrases from a set [15]. In summary, the design was: 12 participants × 5 conditions × 15 English phrases = 900 phrases, in total.

Keyboard Backgrounds/Themes

For the color conditions, we selected two colors from the two sides of a color display spectrum: orange (RGB #FF6F00) and teal (RGB #58C7D1). Orange is a *warm* color that can strain human eyes and often uncomfortable to look at for an extended period time, while teal is a *cool* color that is pleasing to human eyes [10].

We added two image conditions since some users may prefer using image backgrounds instead of colors. For the preselected image condition, we randomly picked an image from the 12 available landscape pictures in Gboard. In the user-selected condition, participants selected an image from the phone's photo gallery. This condition was included with the consideration that some users may prefer using personal pictures as backgrounds, such as photographs of their significant others, children, pets, etc. Since we could not request access to their personal images due to privacy concerns, we tried to replicate this scenario by creating a custom image gallery.

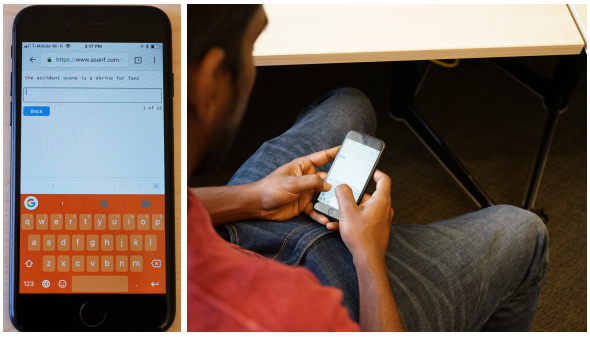

Figure 1. The device and the application used in the study (left); a volunteer participating in the study (right).

For this, first we identified the most popular image categories for desktop, tablet, or smartphone backgrounds in the Internet (e.g., [6,19]). Pets, nature, wildlife, motor vehicles, and skyscrapers were found to be the most popular categories. Then, we downloaded 5 images for each category, resulting in 25 images in the gallery. Figure 2 shows some of the keyboard backgrounds used in the study.

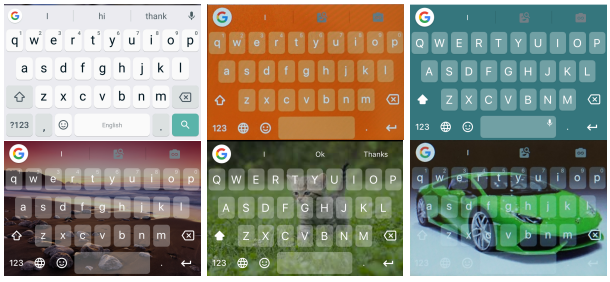

Figure 2. From top left, the default, warm color, cool color, preselected image, and two user-selected image backgrounds.

Procedure

First, we demonstrated Gboard's theme selection feature to all participant and explained the study procedure. Then, we collected their consents and demographics. The study started shortly after that, where each participant entered fifteen phrases in each condition. WebTEM [3] displayed one phrase at a time. Participants had to transcribe each phrase and press the *Return* key to see the next phrase. They were instructed to read and memorize the phrase and then transcribe it as fast and accurate as possible. Error correction was recommended [5]. There was no practice session since all participants were experienced virtual Qwerty users. To increase the external

validity of the study, we allowed participants to enter text in their usual posture and position. Interestingly, all participants held the device with both hands in portrait position, and entered text using the thumbs (Figure 1).

All participants started with the default (baseline) condition, then the other conditions in a counterbalanced order. During the user-selected condition, participants picked an image of their liking from the gallery (included 25 images, see above). There were mandatory short breaks (~one minute) between the conditions. However, participants could request for additional breaks, when necessary. Upon the completion of the study, participants were asked to fill out a short questionnaire, where they could rate the examined keyboards background on a 7-point Likert scale.

Results

The study lasted for about 30 minutes. It recorded the commonly used Words per Minute (WPM), Error Rate (ER), and Keystroke per Characters (KSPC) metrics [5]. We used a repeated-measures ANOVA to analyze all quantitative data. To investigate the effects of different "types" of background, we grouped the conditions into the default, color (warm and cool), and image (preselected and self-selected) categories. In the self-selected background condition, 58% participants picked wildlife, 17% motor vehicles, 17% pets, and the reaming 8% nature. We did not analyze the data for background image categories due to insufficient data.

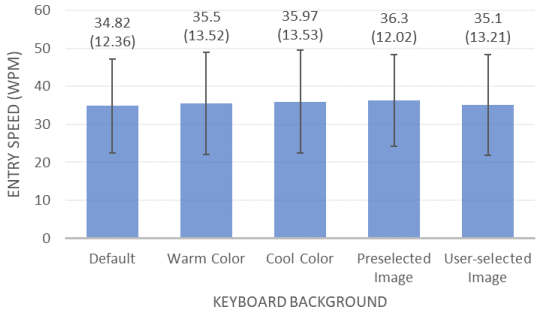

Figure 3. Average entry speed per condition. The values inside the brackets and the error bars represent standard deviations.

Entry Speed (WPM)

An ANOVA failed to identify a significant effect of condition on entry speed ($F_{4,11} = 0.53$, $p > .05$). Figure 3 illustrates average entry speed for all conditions. There was also no significant effect of background type ($F_{2,11} = 0.39$, $p > .05$). Entry speed for default, color, and image were 34.82 (SD = 12.36), 35.74 (SD = 13.52), and 35.7 (SD = 12.64) WPM, respectively.

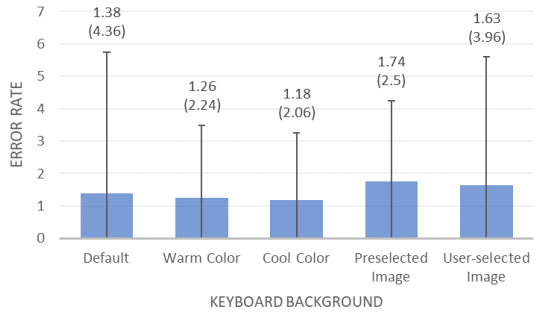

Figure 4. Average error rate per condition. The values inside the brackets and the error bars represent standard deviations.

Error Rate (ER)

An ANOVA failed to identify a significant effect of condition on error rate ($F_{4,11} = 0.82$, $p > .05$). Figure 4 illustrates average error rate for all conditions. There was also no significant effect of background type ($F_{2,11} = 2.4$, $p = .10$). Average error rate for default, color, and image were 1.35% (SD = 4.26), 1.22% (SD = 2.15), and 1.66% (SD = 3.01), respectively.

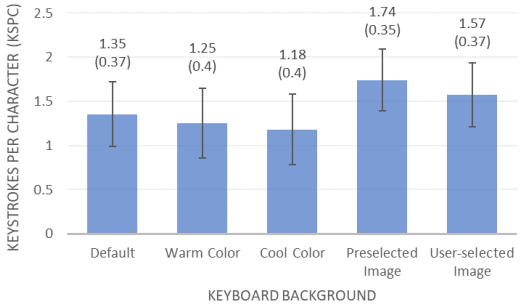

Figure 5. Average Keystrokes per Character (KSPC) for all conditions. The values inside the brackets and the error bars represent standard deviations.

Keystrokes per Character (KSPC)

An ANOVA failed to identify a significant effect of condition on KSPC ($F_{4,11} = 0.84$, $p > .05$). Figure 5 illustrates average KSPC for all conditions. There was also no significant effect of background type ($F_{2,11} = 1.16$, $p = .30$). KSPC for default, color, and image were 1.35 (SD = 0.37), 1.22 (SD = 0.4), and 1.66 (SD = 0.36), respectively.

User Feedback

In the study, participants responded to questions about speed, accuracy, and interference for the color and image conditions in comparison with the default condition on a 7-point Likert scale. We used a Friedman Test to analyze user responses.

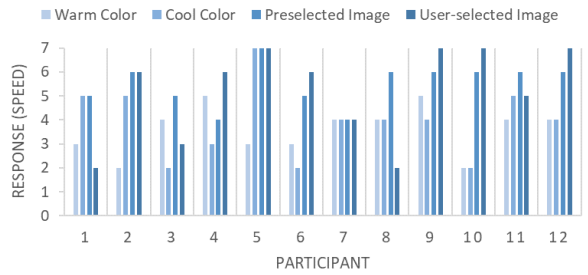

Figure 6. Participant responses to whether keyboard background affected their text entry speed. Here, *1* to *7* represent *Definitely Disagree* to *Definitely Agree*.

Perceived Speed

A Friedman Test identified a significant effect of condition on perceived speed ($\chi^2(3) = 12.03$, $p < .01$). Figure 6 shows user responses to the question whether keyboard background affected their entry speed, where one can see that most participants responded that the image conditions reduced their entry speed. About 83% ($N = 10$) and 67% ($N = 8$) participants felt that the preselected and user-selected image background, respectively, affected their speed. They were mostly neutral about the other conditions (Table 1).

Perceived Accuracy

A Friedman Test identified a significant effect of condition on perceived accuracy ($\chi^2(3) = 9.61$, $p < .05$). Figure 7 shows user responses to the question whether keyboard background affected their accuracy, where one can see (similar to entry speed) most participants responded that the image conditions reduced their accuracy. About 83% ($N = 10$) and 58% ($N = 7$) participants felt that the preselected and user-selected image background, respectively, affected their accuracy. They were mostly neutral about the other conditions (Table 1).

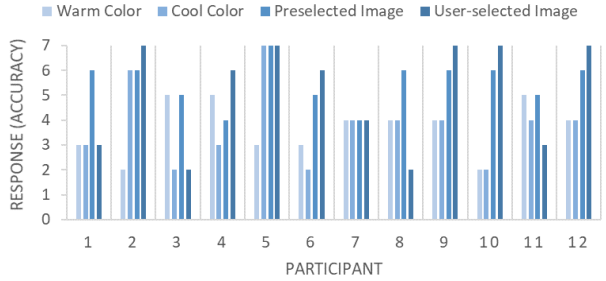

Figure 7. Participant responses to whether keyboard background affected their text entry accuracy. Here, *1* to *7* represent *Definitely Disagree* to *Definitely Agree*.

Interference

A Friedman Test identified a significant effect of condition on interference ($\chi^2(3) = 13.82$, $p < .05$). Figure 8 illustrates user responses to the question whether keyboard background swayed their attention away from text entry, where it is apparent that most participants did not find the backgrounds destructing. Over 67% ($N = 8$) participants either disagreed or were neutral about this.

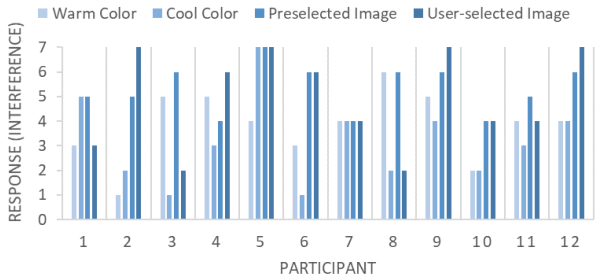

Figure 8. Participant responses to whether keyboard background interfered with the task of text entry. Here, *1* to *7* represent *Definitely Disagree* to *Definitely Agree*.

DISCUSSION

The results do not support acceptance of H_1. Statistical tests revealed that there was no significant effect of keyboard background on either text entry speed, accuracy, or keystrokes per character. It is apparent in Figure 3–5 that all backgrounds yielded comparable speed, accuracy, and KSPC.

Interestingly, the results support acceptance of H_2. Statistical tests identified significant effect of keyboard background on both perceived speed and accuracy. Table 1 presents participant responses to the questions about whether they felt that the examined backgrounds affected their text entry speed or accuracy, from where it is clear that substantially more participants felt that image backgrounds compromised their text entry speed and accuracy, although in reality, that was not the case (Figure 3–5). They were mostly neutral about the color themes, i.e., did not feel that they interfered with text entry.

Background (Theme)	Perceived Speed			Perceived Accuracy		
	Agree	Disagree	Neutral	Agree	Disagree	Neutral
Warm Color	17%	**42%**	**42%**	25%	**33%**	**42%**
Cool Color	**33%**	**33%**	**33%**	17%	**42%**	**42%**
Preselected Image	**83%**	17%	0%	**83%**	17%	0%
User-selected Image	**67%**	8%	25%	**58%**	25%	17%

Table 1. Percentage of users agreed, disagreed, or were neutral about the effects of keyboard backgrounds on speed and accuracy. The numbers do not always add up to 100% since they are rounded to the nearest integers.

The study used the same keyboard in all conditions, but with diffident themes, yet participants had strong opinion about some of the conditions. This is particularly interesting since it suggest that it is possible to influence user impression of a keyboard's performance simply by changing its theme.

Recommendations

Based on the findings, we caution researchers against using radically different keyboard themes in different conditions. While it may be acceptable in studies interested only in quantitative measures, studies collecting user opinion must make sure that all examined keyboards look more or less the same since user preference for a theme could introduce a confounding variable. Further research is necessary to identify the factors that influence user opinion. Keyboard developers must also be careful in tweaking the default or designing new themes. Instead of making sudden and drastic changes,

we recommend gradually switching to new themes by introducing subtle changes. We also recommend enabling the user to select and edit themes through a theme library.

CONCLUSION AND FUTURE WORK

In this paper, we investigated the effects of different types of keyboard backgrounds (themes) on actual and perceived text entry performance, in terms of speed and accuracy. We conducted a user study that compared two color and two image backgrounds with the default Google Android keyboard. Results revealed that there was no significant effect of keyboard background on actual performance, yet, a significant effect on perceived performance was identified. Most participants felt that image backgrounds, regardless of whether they were pre or self-selected, affected their entry speed and accuracy. This indicated towards the possibility that users' text entry experience can be influenced by using effective keyboard themes. Based on these findings, we made design recommendations for text entry researchers and keyboard developers.

In the future, we will extend this work to various properties of image (apply clutter metrics), different experience levels (experts and novices), and investigate whether there is a long-term effect of keyboard themes on performance. We will also explore the possibility of improving text entry experience using keyboards that can identify and adapt to the user's mood (or the current state of mind) by changing themes to provide them with a pleasant text entry experience.

REFERENCES

1. Toshiaki Aiyoshizawa and Takashi Komuro. 2017. Comparative study on text entry methods for mobile devices with a hover function. In *Proceedings of the 16th International Conference on Mobile and Ubiquitous Multimedia - MUM '17*, 355–361. https://doi.org/10.1145/3152832.3156614
2. Ahmed Sabbir Arif, Benedikt Iltisberger, and Wolfgang Stuerzlinger. 2011. Extending mobile user ambient awareness for nomadic text entry. In *Proceedings of the 23rd Australian Computer-Human Interaction - OZCHI '11*, 21–30. https://doi.org/10.1145/2071536.2071539
3. Ahmed Sabbir Arif and Ali Mazalek. 2016. WebTEM: A Web application to record text entry metrics. In *Proceedings of the 2016 ACM on Interactive Surfaces and Spaces - ISS '16*, 415–420. https://doi.org/10.1145/2992154.2996791
4. Ahmed Sabbir Arif, Michel Pahud, Ken Hinckley, and Bill Buxton. 2014. Experimental study of stroke shortcuts for a touchscreen keyboard with gesture-redundant keys removed. In *Proceedings of the 2014 Graphics Interface Conference - GI '14*, 43–50. https://doi.org/10.20380/GI2014.06
5. Ahmed Sabbir Arif and Wolfgang Stuerzlinger. 2009. Analysis of text entry performance metrics. In *Proceedings of the IEEE Toronto International Conference - Science and Technology for Humanity - TIC-STH '09*, 100–105. https://doi.org/10.1109/TIC-STH.2009.5444533
6. Gary Chan. Most Popular Unsplash Desktop Wallpapers. *Unsplash*. Retrieved June 1, 2018 from https://unsplash.com/collections/1041097/most-popular-unsplash-desktop-wallpapers
7. Andrew J. Elliot and Markus A. Maier. 2014. Color psychology: effects of perceiving color on psychological functioning in humans. *Annual Review of Psychology* 65, 1: 95–120. https://doi.org/10.1146/annurev-psych-010213-115035
8. Beverly L Harrison, Hiroshi Ishii, Kim J. Vicente, and William A. S. Buxton. 1995. Transparent layered user interfaces: An evaluation of a display design to enhance focused and divided attention. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '95*, 317–324. https://doi.org/10.1145/223904.223945
9. Takeshi Hatta, Hirotaka Yoshida, Ayako Kawakami, and Masahiko Okamoto. 2002. Color of computer display frame in work performance, mood, and physiological response. *Perceptual and Motor Skills* 94, 1: 39–46. https://doi.org/10.2466/pms.2002.94.1.39
10. Paul Kay and Luisa Maffi. 1999. Color appearance and the emergence and evolution of basic color lexicons. *American Anthropologist* 101, 4: 743–760. https://doi.org/10.1525/aa.1999.101.4.743
11. Jeong Ho Kim, Lovenoor S. Aulck, Ornwipa Thamsuwan, Michael C. Bartha, Christy A. Harper, and Peter W. Johnson. 2013. The effects of touch screen virtual keyboard key sizes on typing performance, typing biomechanics and muscle activity. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 8026 LNCS, PART 2: 239–244. https://doi.org/10.1007/978-3-642-39182-8_28
12. Jeong Ho Kim, Lovenoor 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. https://doi.org/10.1177/0018720814531784
13. Jianwei Lai, Dongsong Zhang, and Sen Wang. 2018. A thumb stroke-based virtual keyboard for sight-free text entry on touch-screen mobile phones. In *Proceedings of the 51st Hawaii International Conference on System Sciences - HICSS-51*, 293–302.
14. Joung-Youn Lee. 2015. Personal computer wallpaper user segmentation based on Sasang typology. *Integrative Medicine Research* 4, 1: 34–40. https://doi.org/10.1016/j.imr.2015.01.001
15. I. Scott MacKenzie and R. William Soukoreff. 2003.

Phrase sets for evaluating text entry techniques. In *CHI '03 extended abstracts on Human factors in computer systems - CHI '03*, 754. https://doi.org/10.1145/765968.765971

16. Aaron Marcus. 1997. Graphical User Interfaces. In *Handbook of Human-Computer Interaction*. Elsevier, 423–440. https://doi.org/10.1016/B978-044481862-1.50085-6
17. Gerald M. Murch. 1984. Physiological principles for the effective use of color. *IEEE Computer Graphics and Applications* 4, 11: 48–55. https://doi.org/10.1109/MCG.1984.6429356
18. C. M. Pancake. 1995. Principles of color use for software developers. In *Tutorial M1 from Supercomputing 1995*.
19. Jessica Probus. 2014. 26 Desktop Backgrounds That Will Make You Not Hate Working. *BuzzFeed*. Retrieved June 1, 2018 from https://www.buzzfeed.com/jessicaprobus/26-remarkably-soothing-desktop-backgrounds
20. Martin Schedlbauer. 2007. Effects of key size and spacing on the completion time and accuracy of input tasks on soft keypads using trackball and touch input. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 51, 5: 429–433. https://doi.org/10.1177/154193120705100501
21. 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. https://doi.org/10.1080/01449299308924362
22. Mary E. Sesto, Curtis B. Irwin, Karen B. Chen, Amrish O. Chourasia, and Douglas A. Wiegmann. 2012. Effect of touch screen button size and spacing on touch characteristics of users with and without disabilities. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 54, 3: 425–436. https://doi.org/10.1177/0018720811433831
23. Steve N H Tsang, H S Chan, and K Chen. 2013. A study on touch screen numeric keypads: Effects of key size and key layout. In *Proceedings of the International MultiConference of Engineers and Computer Scientists - IMECS '13*, 13–16. Retrieved from http://www.iaeng.org/publication/IMECS2013/IMECS2013_pp1006-1009.pdf
24. Keith Vertanen, Crystal Fletcher, Dylan Gaines, Jacob Gould, and Per Ola Kristensson. 2018. The impact of word, multiple word, and sentence input on virtual keyboard decoding performance. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–12. https://doi.org/10.1145/3173574.3174200
25. Peggy Wright, Diane Mosser-Wooley, and Bruce Wooley. 1997. Techniques & tools for using color in computer interface design. *Crossroads* 3, 3: 3–6. https://doi.org/10.1145/270974.270976
26. FancyKey. Retrieved August 24, 2018 from https://itunes.apple.com/us/app/fancykey-keyboard-themes/id1073000685
27. RainbowKey. Retrieved August 24, 2018 from https://itunes.apple.com/us/app/rainbowkey-color-keyboard-themes-fonts-gif/id915637540
28. Best Color Keyboard. Retrieved August 24, 2018 from https://play.google.com/store/apps/details?id=com.jb.gokeyboard.theme.tkbestcolorkeyboard
29. Change Color of Keypad. Retrieved August 24, 2018 from https://play.google.com/store/apps/details?id=com.jb.gokeyboard.theme.twchangecolorofkeypad
30. Pastel Color Keyboard. Retrieved August 24, 2018 from https://play.google.com/store/apps/details?id=com.ikeyboard.theme.watercolor2017
31. Gboard - The Google Keyboard. Retrieved June 1, 2018 from https://play.google.com/store/apps/details?id=com.google.android.inputmethod.latin