Summer 2015

Comparing Taxi Clearance Input Layouts for Advancements in Flight Deck Automation for Surface Operations

Lara W.S. Cheng
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

COMPARING TAXI CLEARANCE INPUT LAYOUTS FOR ADVANCEMENTS IN
FLIGHT DECK AUTOMATION FOR SURFACE OPERATIONS

A Thesis

Presented to

The Faculty of the Graduate Program in Human Factors and Ergonomics

San José State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science

by

Lara W.S. Cheng

August 2015

The Designated Thesis Committee Approves the Thesis Titled

COMPARING TAXI CLEARANCE INPUT LAYOUTS FOR ADVANCEMENTS IN
FLIGHT DECK AUTOMATION FOR SURFACE OPERATIONS

by

Lara W.S. Cheng

APPROVED FOR THE DEPARTMENT OF
INDUSTRIAL AND SYSTEMS ENGINEERING

SAN JOSÉ STATE UNIVERSITY

August 2015

Dr. Sean Laraway	Department of Psychology
Dr. Kevin Jordan	Department of Psychology
Dr. Becky Hooey	San José State University Research Foundation

ABSTRACT

COMPARING TAXI CLEARANCE INPUT LAYOUTS FOR ADVANCEMENTS IN
FLIGHT DECK AUTOMATION FOR SURFACE OPERATIONS

By Lara W.S. Cheng

Airport moving maps (AMMs) have been shown to decrease navigation errors, increase taxiing speed, and reduce workload when they depict airport layout, current aircraft position, and the cleared taxi route. However, current technologies are limited in their ability to depict the cleared taxi route from ATC to the flight deck. This study examined methods by which pilots can input ATC-issued taxi clearances to support taxi route depictions on the AMM. Sixteen general aviation (GA) pilots used a touchscreen monitor to input taxi clearances using two input layouts, softkeys and QWERTY, each with and without feedforward (graying out invalid inputs). QWERTY yielded more taxi route input errors than the softkeys layout. The presence of feedforward did not produce fewer taxi route input errors than in the non-feedforward condition. The QWERTY layout did reduce taxi clearance input times relative to the softkeys layout, but when feedforward was present this effect was observed only for the longer, 6-segment taxi clearances. With the softkeys layout, feedforward reduced input times compared to non-feedforward but only for the 4-segment clearances. Feedforward did not support faster taxi clearance input times for the QWERTY layout. Based on the results and analyses of the present study, it is concluded that for taxi clearance inputs, (1) QWERTY remain the standard for alphanumeric inputs, and (2) feedforward be investigated further, with a focus on participant preference and performance of black-gray contrast of keys.

ACKNOWLEDGEMENTS

First and foremost, I'd like to thank my family. In particular, Mommy, Daddy and Nadia, you have served as my biggest supporters and inspirations, always. Thank you for always checking in on me, keeping me close, and always being reliable. Your unwavering encouragement and love have allowed me to become the best I can be, and I will never stop striving for improvement. I'd also like to thank the rest of my family for being supportive of my endeavors my entire life. You continue to give me confidence in the person I am, the decisions I make, and more importantly, the lessons I have learned and applied to my life.

To my thesis committee and mentors, Drs. Sean Laraway, Kevin Jordan and Becky Hooey, I thank you for your tireless efforts to guiding me to become the best professional I can be, while staying true to myself. Dr. Jordan, thank you for your support over the past several years, and for always believing in me. Your encouragement has allowed me to grow as an individual and given me confidence to pursue my dreams. You have taught me to accept my flaws, but also to persevere in all aspects of my life. Dr. Laraway, thank you for supporting me as a student, professional, and individual. I'm so grateful that I have had the opportunity to pursue my goals and interests, and to know that I can always turn to you. Dr. Hooey, thank you for always encouraging me and teaching me the skills that will help me succeed in my career. All the time and effort you have put into guiding me in the lab and on my thesis has helped build my confidence as a researcher; and thanks to you, I know that I will always continue learning and improving

as a professional. Finally, thank you all for being so incredibly patient, especially in my times of struggle.

I'd also like to thank Dr. Dave Foyle for giving me such immense opportunities over the past few years, and for having me as a part of the Human-Centered Systems Laboratory. I'd also like to thank Debi Bakowski, Glenn Meyer and Cynthia Wolter for their support in the research group, and for helping make my adjustment into graduate life less daunting.

A special thanks goes out to my mentors Sandy Lozito and Savita Verma, my first two mentors at NASA Ames. I will always appreciate the tasks you trusted me to carry out and for giving me the most phenomenal first internship experiences. Thank you for helping kickstart my career!

A huge thank you goes out to all others who helped me pull my thesis together (there are several more than I can mention here). Conrad Gabriel Jr., thank you for finding me participants! Thanks to you stepping up and reaching out to people for me, I was able to run my thesis. Keith Chin, buddy since middle school and one of my best friends since post-college, I don't even know how to thank you for getting my experiment to work! Thank you for all the hours you spent helping me build my thesis conditions, for always being patient, and for always being up for the high school reminiscing. To Bruno from Paradigm, thank you for all the help you provided throughout the thesis process—I could not have finished this thesis without your help!

My dearest, Jake, thank you for always being there for me, especially when I need it most. You are so incredibly patient, caring, honest and loving—I could never hope for a better person to be with. Finally, I'd like to thank all my friends who have supported and encouraged me as a professional and individual over the years (YOU KNOW WHO YOU ARE). You always make me improve as an individual, you keep life interesting, and you can always put a smile on my face.

Table of Contents

List of Figures

List of Tables

Introduction

In current-day surface operations, an air traffic controller (ATC) issues a taxi clearance to pilots via radio. The taxi clearance typically consists of the aircraft call sign, a destination (runway or concourse), and a series of taxiways to follow to get to the destination. At all airports, taxiways are labeled using combinations of alphabetical and numerical characters such as A, B2, H, and the controller uses the International Civil Aviation Organization (ICAO) phonetic alphabet (e.g., Alpha for A) to prevent misunderstandings. An example of a typical taxi clearance is "Airline 123, Taxi to Runway 17R via Alpha, Bravo, Charlie." Pilots either write down the clearance on a scratch pad or memorize the clearance. They are required to read the clearance back to the ATC. Pilots typically taxi using a paper chart and airport signage for navigation.

As discussed by Schönefeld and Möller (2012), if it is the case that pilots are not familiar with the airports into or out of which they are flying, pilot errors can occur during the taxi phase such as a wrong turns, increased heads-down time required to consult navigation charts, and increased workload. Across two high fidelity pilot-in-the-loop simulations, Hooey and Foyle (2006) found that pilots made a navigation error on approximately 17% of trials conducted in low visibility conditions at a complex airport (Chicago O'Hare International Airport). At best, these navigation errors reduce the efficiency of airport operations. At worst, they have serious safety consequences such as runway incursions.

Airport moving maps (AMMs) have been proposed as a way to decrease navigation errors, increase taxi efficiency, and reduce workload (Hooey & Foyle, 2006; Theunissen, Roefs, Koeners, & Bleeker, 2007). The recent introduction of electronic flight bags, portable GPS units, and tablet computers enables the widespread adoption of these AMMs on the flight deck. Currently, AMMs show the layout of the airport concourses, taxiways, and runways. GPS-equipped units allow for real-time updates of the aircraft's current position. To date, none shows the cleared-to-taxi route as issued by ATC due to the unavailability of datacomm or other means of electronically transmitting clearances from ATC to the flight deck. However, empirical evidence suggests that displaying the cleared taxi route reduces taxi errors and enables conformance monitoring (Hooey & Foyle, 2006; Theunissen et al., 2007).

In previous research, Theunissen et al. (2007) tested pilots' ability to input taxi clearances manually to support displaying the route on flight-deck AMMs. They emphasized that taxi route input methods should allow pilots to enter the routes accurately, quickly and with such low workload that the pilots are able to complete the task while receiving the radio ATC clearance so as to integrate seamlessly into current-day operations. This also serves as a means for error checking. Reading back the input clearance to ATC from the AMM display serves to close the integrity loop by allowing both ATC and pilots an additional opportunity to catch any potential errors (Theunissen et al., 2007).

Theunissen et al. (2007) tested an input method that presented virtual keys on a touchscreen overlaying the AMM. They used 20 touchscreen keys with fixed labeling to

represent the taxiways at the test airport. They noted that this limited number of keys might not be sufficient for complex airports, which may require an added page scrolling (or similar) function to accommodate the taxiway labels. Theunissen et al. (2007) also tested a context-sensitive input function in which only valid taxiways connected to the last entered leg are available for the user to select.

The present study compared various touchscreen taxi route input layouts in order to contribute to the previous studies' findings of various aids to assist pilots during the taxi phase. This study evaluated performance and preference of a current-day baseline condition and four input layouts that combine the concepts of Theunissen et al.'s (2007) virtual keys and context-sensitive input functions as well as a QWERTY layout. The QWERTY layout is proposed as a simple input layout with which pilots are familiar and which may alleviate the issues identified by Theunissen et al.'s virtual key concept – specifically eliminating the clutter and complexity of too many virtual keys and potential page-scrolling requirements at complex airports.

Taxi Route Input Devices

Although the long-term intent is to transmit taxi clearances to the flight deck directly from ATC via datacomm (Truitt & Muldoon, 2010), this idea is still very futuristic, and the infrastructure is not in place to support it. Various options exist that allow pilots to manually enter the taxi route, two of which are voice input and touchscreens. Theunissen, Roefs, Moncur, and Jinkins (2006) suggested one voice input method that incorporates speech recognition. However, Roefs, Theunissen, and Koeners

(2007) subsequently found that manual input of taxi clearances yielded better

performance (speed and efficiency) and workload ratings than voice input.

Touchscreens allow for much more information to be condensed into a small

interface. This approach has been adopted by many car manufacturers in order to

implement many features that serve as secondary driving tasks such as navigation,

maintaining a safe distance from surrounding vehicles, and hazard detection (Matthews,

Bryant, Webb, & Harbluk, 2001) as well as enhancing the driving experience by allowing

drivers to check route information, pick up and make phone calls, send text messages,

and change the in-vehicle climate, to name a few common examples. However, with

such technological advancements come potentially harmful tradeoffs including decreased

out-the-window situational awareness and overall safety (Regan, Lee, & Young, 2009).

It is for this very reason that some car manufacturers, such as Ford Motor Company, are

eliminating touchscreens in some of their models and switching back to hard dials and

buttons (Howard, 2014). Still, in this study, a touchscreen was used because

touchscreens allow for interface layout flexibility. In addition, the task that was asked of

participants, inputting taxi clearances, would be performed only when the aircraft was

stopped (typically before entering the airport movement area).

The Current Study

This study examined *touchscreen formats* that enable pilots to input taxi routes as

issued from ATC to be able to display the route on the AMM. As Theunissen et al.

(2007) pointed out, a taxi route input method must enable fast, error-free entry of taxi

routes. They note that it is critical that taxi route input methods allow the pilot to enter the taxi route while ATC is issuing the clearance. This allows the pilots to read back the full clearance as entered, without delay, which is critical given the fast-paced environment at busy airports. Further, it allows both the pilot and ATC to error check the route as entered to prevent errors. This provides an extra layer of error checking.

Softkeys are buttons that are context-sensitive. That is, the function of buttons or keys can change depending on the context of the system mode (Kiljander, 2004). Pilots currently use softkeys on the control display unit (CDU) to enter flight plans into the flying model simulator (FMS) (see Figure 1), though instead of a touchscreen the buttons are physical edge keys that line the left and right sides of the display. The label is shown on the display, and changes based on the mode (Midkiff, Hansman, & Reynolds, 2004). To use softkeys, the user must aim a finger at a specific key on the keypad and ensure that the entry was correct by visually crosschecking with the display (Ha, Inkpen, Mandryk, & Whalen, 2006).

Figure 1. Control display unit in a Boeing 737-300.

Given that pilots have experience with softkeys displays for en route flight plan input, the softkeys input layout was selected as one of the taxi route input layouts for this study. For taxi route input, a touchscreen display could use softkeys labeled with airport taxiway names, with the labels changing as relevant for the particular airport. The QWERTY keyboard was selected as the second input layout for the present study because it is listed as the Federal Aviation Administration (FAA) standard (Ahlstrom & Longo, 2003), which refers to the Department of Energy as its source for this feature (Department of Energy, 1992).

Norman (2002) stated that user error should be prevented by not allowing users to make errors, or at least the interface should make undoing an error a simple task. Feedforward is an interface design feature that is meant to reduce user errors (Norman, 2002), and is a mechanism that guides future interactions via cues (Nielsen, 1995). One example of feedforward used in navigation devices is the graying-out of invalid entries or menu items. The Magellan RoadMate 1212 (Magellan Naviation, Inc., Santa Clara, CA) (Figure 2), a GPS device generally used for navigating while driving, and the Magellan mobile applications incorporate an error-reducing gray-out technique to prevent users from making errors as the user types in a street or city name.

Figure 2. Magellan RoadMate with gray-out function.

Such input layouts for "predictive" or feedforward auto-complete have been studied, where pilots need only enter a few of the taxi segments in order for common routes to be displayed (Theunissen, Roefs, Koeners, & Bleeker, 2008). The present study did not display complete or common route suggestions, but rather only the possible connecting taxiways. "Next possible options" rather than auto-complete was selected as the feedforward method for this study because this might allow for users to be more accurate as they check each segment one by one. Specifically, two touchscreen input layouts were evaluated: softkeys and QWERTY keyboard.

Rationale/Purpose for the Study

The present study contributes to research at NASA in terms of making advancements in NextGen, in that it explores methods to improve performance levels to maximize capacity (Joint Planning and Development Office, 2011), via advancements in flight-deck interfaces and technology. Specifically, the study compared *softkeys vs. QWERTY* and *non-feedforward vs. feedforward.*

Hypotheses

The following hypotheses were tested:

1. Feedforward conditions will yield fewer input errors than non-feedforward conditions will because users' options are limited only to valid inputs, thus *reducing* but not *eliminating* opportunity for error.

2. QWERTY input layout conditions will yield fewer input errors than softkeys conditions because QWERTY keyboards are more common for "everyday" technological devices (e.g., smartphones, touchscreen tablets, computer keyboards).

3. Taxi clearance input time will be faster in the feedforward conditions than in the non-feedforward conditions because in the feedforward conditions, unavailable options are grayed out or unavailable so the user will not need to even consider invalid options, whereas in the non-feedforward conditions, all the options are always available, so users have to determine each of the options as either "correct or incorrect input."

4. Taxi clearance input times will be faster in the QWERTY conditions than in the softkeys conditions because QWERTY keyboards are more common for "everyday" technological devices (e.g., smartphones, touchscreen tablets, computer keyboards).

Method

Participants

The present study was approved by both the NASA Ames Human Research Institutional Review Board (Appendix A) and the Human Subjects Institutional Review Board at San José State University (Appendix B). An email announcing the study was sent to general aviation (GA) pilots who are in the San José State University Test Subject Recruitment Office's (TSRO) participant database. A number of respondents participated in the study, and the remaining participants were recruited via snowball sampling. Participants were compensated financially for their time based on current rates for GA pilots. The participants in this study were 16 (15 males, 1 female) GA pilots with experience using an avionics suite, such as the Garmin G1000 or G430. All participants had a valid GA license at the time of testing, had flown within 15 days of the start of the study, and signed a consent form upon arrival to the study (which also included a short description of the study) (Appendix C). The mean age of the participants was 29.75 years old.

Apparatus/Devices

A 21.5" HP Compaq L2105tm widescreen touchscreen monitor was used as the input device, on which participants entered the taxi clearances using the four different conditions. Each of the interfaces was interactive in that the inputs were typed on the screen as the participants tapped on the corresponding buttons. The touchscreen monitor was also used for post-condition workload questionnaires that were based on the NASA task load index (TLX) (Hart & Staveland, 1988) (Appendix D). The touchscreen monitor

was connected to a 13.3" MacBook Air via an Apple Mini Display Port to VGA Display Adapter. On the MacBook Air, the researcher launched each of the four conditions for the participant. The four conditions were designed using Paradigm, which was the program used to build the experiment and run the study, a product of Perception Research Systems Incorporated, Version 2.3.0.29 [x64] (see Appendix E for a detailed explanation for choosing this software). Participants also used the MacBook Air to fill out and type responses to the demographics questionnaire (Appendix F). An Asus VivoBook 13.3" touchscreen laptop was also used in the experiment, from which the researcher played the taxi clearance audio clips for the participant. This laptop was connected to a Jawbone JAMBOX Wireless Bluetooth Speaker (5.94" × 1.57" × 2.24"; 12 ounces). Both the MacBook Air and the Asus VivoBook were running Windows 8 (Microsoft, Inc., Redmond, WA). Figure 3 shows how the devices were configured during the study.

The alphanumeric keyboard format selected for the present study is a QWERTY layout, as it has been determined by FAA to be the standard (Ahlstrom & Longo, 2003). The keyboard format met, or exceeded, the following FAA requirements:

- 9.1.7 Key size. The minimum horizontal surface width for a key on a typing keyboard should be 12 mm. [Source: American National Standard Institute (ANSI), 1988]

- 9.1.8 Horizontal spacing of keys. Horizontal centerline distances should be between 18-19 mm. [Source: ANSI, 1988]

- 9.1.9 Vertical spacing of keys. Vertical centerline distances should be between 18-21 mm. (p. 9-5)

The four conditions used in this study are shown in Figure 4. For this study, numbers were used in combination with letters (F2, G8, H9, K5) as their own taxiways. On the QWERTY keyboard, each of those pairs has the same distance between letter and number; the purpose of this was to eliminate the variable of distance between keys.
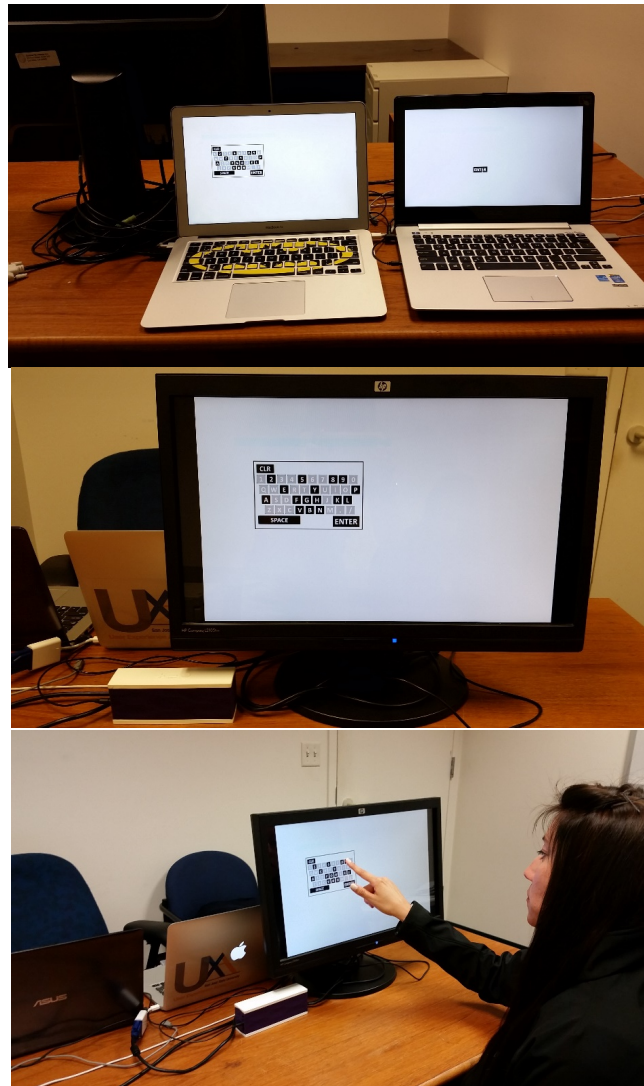
Figure 3. Study setup. Top: Researcher's point of view with MacBook Air on the left and Asus VivoBook on the right. Middle: Participant's point of view with touchscreen monitor directly in front and JAMBOX on bottom left side of monitor. Bottom: Overall setup of devices.
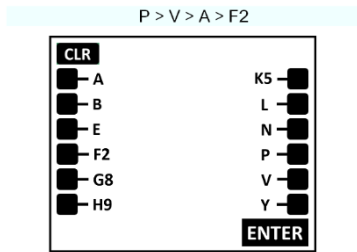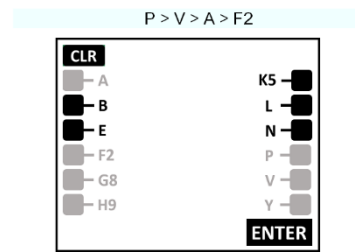
softkeys without feedforward   softkeys with feedforward
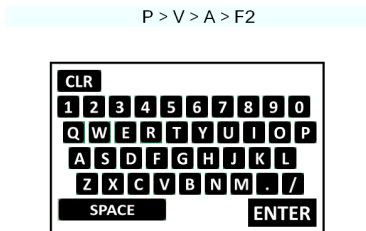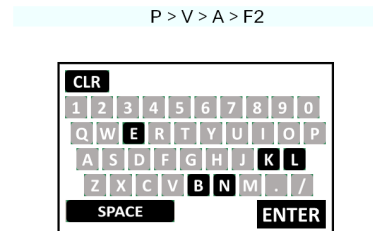
QWERTY without feedforward   QWERTY with feedforward

Figure 4. Four input conditions used in the present study. The following example represents what each of the four conditions would look like: (1) the assigned taxi route is P > V > A > F2 > K and (2) the only taxiways that can be branched from F2 are B, E, K (for QWERTY, and K5 in the softkeys layout), L and N.

Laptop synchronization. Because two laptops were used in this experiment and millisecond timestamp accuracy was required, it was necessary to synchronize the laptops' times. The method selected was that of the National Institution of Standards and Technology's method for synchronizing via the Internet (National Institute of Standards and Technology, 2013). Laptop synchronization was performed before each condition was run.

Materials

Upon arriving to the study, participants were handed a paper with a brief description of the study, as well as a consent form (Appendix C) to read and fill out. The

12

study also included a number of questionnaires: post-condition workload questionnaires (completed on the MacBook Air, see Appendix D), a post-study questionnaire (filled out on paper, see Appendix G), and a demographics questionnaire (completed on the MacBook Air, see Appendix F). The researcher followed a paper script (see Appendix H) throughout the study, from which instructions and participant expectations were recited to the participants. There was also a checklist (see Appendix I) that the researcher used between trials and between participants to ensure that each step of the study was consistently set up for all participants. The post-condition workload questionnaire was based on the NASA-TLX, and participants filled it out on the touchscreen monitor after each of the conditions. The post-condition workload questionnaire was comprised of the following seven questions, all of which had a 7-point, Likert-type scale (1 = Very Low, 7 = Very High):

1. Overall workload: How would you rate the overall workload in this condition?
2. Mental Demand: How mentally demanding was the task?
3. Physical Demand: How physically demanding was the task?
4. Temporal Demand: How hurried or rushed was the pace of the task?
5. Performance: How successful were you in accomplishing what you were asked to do?
6. Effort: How hard did you have to work to accomplish your level of performance?
7. Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you?

The demographics questionnaire collected information on the participants' flying experience, experience with technology (such as whether or not they own a smartphone and/or tablet), as well as their age and gender. At the end of the study, participants completed a post-study questionnaire in which they ranked the conditions in order of preference, rated ease of use as well as safety, and compared each of the four touchscreen conditions with the written condition. The written condition was not analyzed in this study because the purpose was to compare input layouts that might support taxi route displays on the flight deck. The reason the written condition was conducted was to allow for potential further analyses in the future, in which input error and input times via flight-deck devices could be compared to current-day writing times and errors for taxi clearances. Aspects such as safety and situation awareness, for example, could be impacted by technological advances, and it may be determined that ultimately, current-day methods (writing) are preferred, safer, or more efficient, to name a few possible benefits.

The taxi clearance audio recordings were read-aloud voice recordings created by the researcher using the military phonetic alphabet (e.g., "taxi via Bravo, Foxtrot, Alpha). For purposes of accuracy (in terms of being able to calculate the difference between the audio start time and the participants' first input times in each trial), the audio clips were edited using the program iMovie. They were clipped at the beginning of the first sound peak so that there was little to no "white noise."

Procedure

For each trial, the researcher played the taxi clearance audio from the Asus VivoBook, and the participant simply input the taxi clearance into the condition display on the touchscreen. With each individual keystroke, the corresponding input was added to the response string above the keys (either the softkeys or the QWERTY keyboard. See Figure 4).

At the beginning of each condition, the researcher described how the input display would look as the participant typed the taxi clearances using the touchscreen. The researcher then demonstrated two trials by playing the taxi clearance audio from the Asus VivoBook and inputting the taxi clearance on the display, but used the cursor on the MacBook Air to click the keys rather than the participant's touchscreen. The participant's monitor mirrored the researcher's MacBook Air for this study, so he or she was able to see the inputs being made on the researcher's side. In the second trial that the researcher demonstrated, an input "error" was intentionally made in order to demonstrate how to clear an incorrect entry and retype the correct route. After the two demonstration trials by the researcher, the participant had the opportunity to ask any questions about the condition before completing four practice trials on his or her own. After having questions answered (if applicable), the participant completed four practice trials and was requested to make an intentional "error" on the second practice trial, to practice the "Clear" function. After the four practice trials, the participant was again asked if he or she had any questions before starting the condition's "experimental" trials. After having questions answered (if applicable), the participant began the set of trials for the given

condition. After entering each taxi clearance, the participant pressed the "Enter" button, at which point the trial was complete and there was a 3-second blank white screen on the participant's monitor. This 3-second blank white screen served as the "pause" between trials. If the participant noticed that the taxi clearance input did not match the taxi clearance audio, then he or she would have to press "Clear" and re-enter the taxi clearance to the best of his or her ability. There was no "Delete" key for single-character deletions in this study. After pressing "Enter" of the final trial in a condition, the researcher told the participant he or she had completed the condition.

For the feedforward conditions, it was explained to participants that black keys were valid inputs (i.e., black alphanumeric characters were taxiways that intersected with the most recently entered taxiway). Gray keys signified taxiways that did not intersect with the most recently entered taxiway. Keys changed between being gray or black with each input.

Design and Analysis

This study used a fully within-subjects factorial design in which all participants completed all trials. The taxi clearances were created by the researcher using a specific set of letters and numbers. In addition, each taxi clearance consisted of exactly one letter-number taxiway. There were 12 taxi clearances with four taxiways, 12 with five taxiways, and 12 with six taxiways. There were 4 practice trials (same order for all conditions) and 36 study trials (randomized once for each of the conditions, prior to the study. See Appendix J). The order of conditions was determined using Latin Square (See Appendix K). The independent variables were: (a) input layouts: entering taxi route

16

using either softkeys or by typing in characters via the QWERTY keypad, and (b) non-feedforward/feedforward feature (e.g., for softkeys input layout, feedforward provided only the available options in black and invalid options in gray from screen to screen, whereas non-feedforward listed all taxiways in black from screen to screen). The 12 taxiways on the display never changed location. In the QWERTY input layout, feedforward faded inputs that are not valid options (e.g., if "A" was pressed, only the taxiways that intersected A were black), whereas the non-feedforward condition did not have the dynamic data gray-out keyboard feature. See Figure 4 for examples of each condition.

Results

The main dependent variables were input errors, input time, workload, Condition Ranking preferences, ease of use, and safety. For input errors and taxi clearance input time, an Input Layout (2) × Non-Feedforward/Feedforward (2) × Taxi Clearance Length (3) ANOVA (alpha = .05) was conducted to compare the four interface styles, specifically whether layouts and/or feedforward had effects. (Taxi clearance length was not a variable for workload, ease of use, and safety because participants were not asked to break down their subjective ratings by taxi clearance length.) Out of 2299 trials, 258 contained errors. The study had a total of 2304 trials, but one trial was omitted due to an outlier in input time. (Any participant mean "input time" for a given condition was deemed an outlier if the mean exceeded 3 standard deviations of the mean for all participants in that condition.) In addition, in five of the trials there were errors on the researcher's part in which a taxi clearance audio clip was started early (i.e., still during

the three-second between-trial pause). For these trials, the participants were told to skip

the input. These dependent variables were analyzed using SPSS Version 20 (IBM, Inc.,

Chicago, IL) and Microsoft Excel (Microsoft, Inc., Redmond, WA).

Input Error

An error was recorded if at least one error was made (e.g., extra key presses,

wrong key presses, missing key presses, or pressing the "Clear" button). If "Clear" was

ever pressed in a trial, the trial was considered to be an "error" trial, regardless of whether

participant re-entered the taxi clearance correctly or not. The reason "cleared-then-

corrected" inputs were counted as errors was that the program did not have the capability

to simply delete a single character. It made more sense, from a research perspective, to

classify an error as any time a trial contained an input other than what was provided by

the taxi clearance audio. The number of trials containing an error was recorded, so if a

participant made more than one error on a single trial, the trial was counted as a single

error. It was hypothesized that feedforward conditions would yield fewer input errors

than non-feedforward conditions because feedforward prevents errors by identifying non-

valid taxiways. Further, it was expected that the QWERTY conditions would yield fewer

input errors than the softkeys conditions because QWERTY keyboards are more common

for "everyday" technological devices (e.g., smartphones, touchscreen tablets, computer

keyboards).

Results for input error across trials. Figure 5 summarizes the data collected for input

error. Figure 6 shows the breakdown of total numbers of errors per participant

18

throughout the study. Table 1 summarizes the results of the 2×2×3 (Input Layout × Non-Feedforward/Feedforward × Taxi Clearance Length) ANOVA.
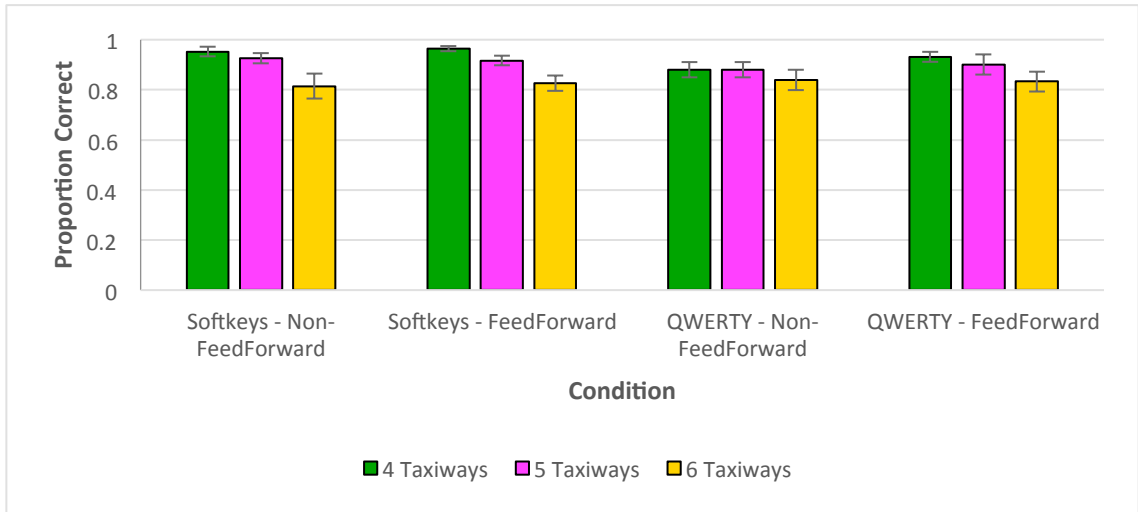
Figure 5. Mean proportion of taxi clearances inputted correctly. Shown with ±1 SE.
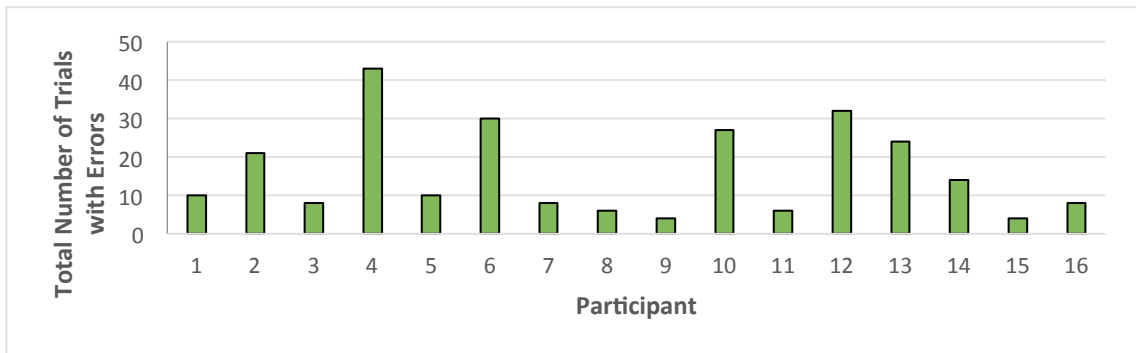
Figure 6. Total errors per participant across all trials.

Table 1

Input Layout × Non-Feedforward/Feedforward × Taxi Clearance Length Repeated Measures ANOVA for Input Errors

Source	df	MS	F	Sig.
(A) Input Layout	1	0.024	2.21	.15
(B) Non-Feedforward/Feedforward	1	0.009	0.82	.38
(C) Taxi Clearance Length	2	0.188	13.44	< .001
A × B (interaction)	1	0.004	0.45	.51
A × C (interaction)	2	0.019	3.46	.044
B × C (interaction)	2	0.004	0.56	.57
A × B × C (interaction)	2	0.004	0.59	.56

There was a significant main effect on taxi clearance length (as seen in Figure 5), $F(2, 30) = 13.44, p < .001$. Paired t-tests showed that while 4- and 5-segment taxi clearances did not have a difference in input error, they both yielded fewer input errors than 6-segment taxi clearances, $t(15) = 4.17, p = .001, d = 0.93$ and $t(15) = 3.62, p = .003, d = 0.72$, respectively. This suggests that 6-segment taxi clearance lengths yield more input errors in general.

There was a significant interaction between input layout (2) and taxi clearance length (3), $F(2, 30) = 3.46, p = .044$. Post-hoc analyses showed that at the 4-segment taxi clearance length, softkeys yielded significantly fewer input errors than QWERTY, $t(15) = 3.10, p = .007, d = 0.79$ (see Figure 7). However, there was no difference for the 5- or 6-segment taxi, $t(15) = 1.43, p = .17, d = 0.35$ and $t(15) = 0.35, p = .54, d = 0.14$, respectively. Contrary to the hypothesis, QWERTY did not yield fewer input errors than softkeys for longer taxi clearances, and in fact yielded more input errors than softkeys for

4-segment taxi clearances. A possible explanation might be that the softkeys layout simply had fewer keys and had much more width between the keys than the QWERTY layout did, though both layouts conformed to the FAA's minimum spacing requirements. The three-way (2×2×3) interaction (Input Layout (2) × Non-Feedforward/Feedforward (2) × Taxi Clearance Length (3)) was not significant, $F(2, 30) = .59$, $p = .56$.
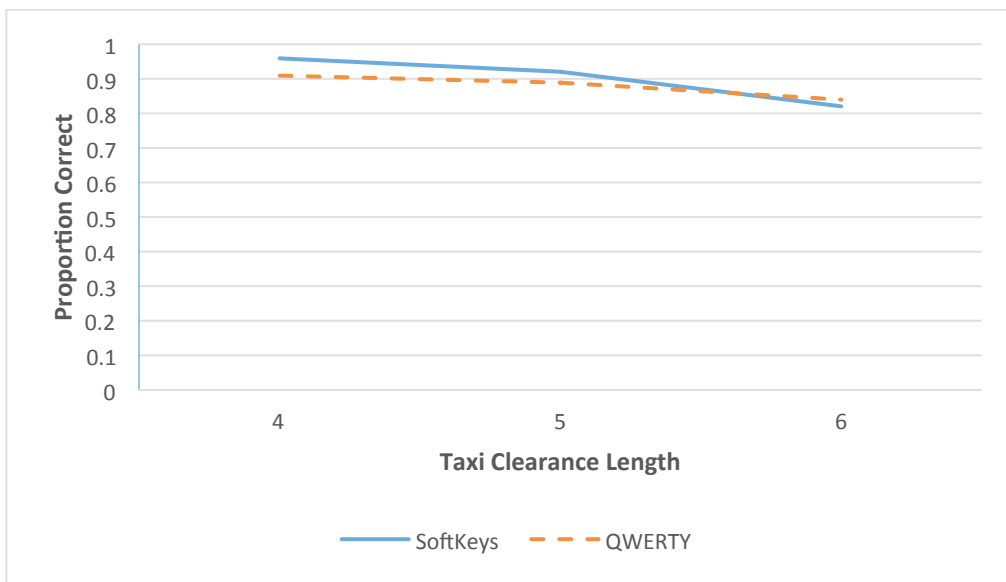
Figure 7. Interaction for input errors: Input Layout (2) × Taxi Clearance Length (3).

Taxi Clearance Input Times

For all conditions, taxi clearance input time was measured from the time that the taxi clearance audio began to the time that the final input was made, prior to pressing "Enter." Only trials that did not contain an error (as defined above) were considered for these analyses. It was hypothesized that taxi clearance input times would be faster in the feedforward conditions than in the non-feedforward conditions because in the

feedforward conditions unavailable options were grayed out reducing the number of options from which to choose. In the non-feedforward conditions, all character options were always black and could be pressed so users had to determine each of the options to be either "correct or incorrect input." It was also hypothesized that taxi clearance input times would be shorter in the QWERTY conditions and in the softkeys conditions because participants have more experience using QWERTY keyboards as they are common on "everyday" technological devices (e.g., smartphones, touchscreen tablets, computer keyboards).

Results for taxi clearance input times. Figure 8 summarizes the data collected for taxi clearance input times. Table 2 summarizes the results of the 2×2×3 (Input Layout × Non-Feedforward/Feedforward × Taxi Clearance Length) ANOVA.

Figure 8. Mean taxi clearance input times. Shown with ±1 SE.

Table 2

Input Layout × Non-Feedforward/Feedforward × Taxi Clearance Length Repeated Measures ANOVA for Input Times

Source	df	MS	F	Sig.
(A) Input Layout	1	7524742	12.84	.003
(B) Non-Feedforward/Feedforward	1	166915	0.73	.41
(C) Taxi Clearance Length	2	52013136	264.24	< .001
A × B (interaction)	1	1784243	2.67	.12
A × C (interaction)	2	706286	3.92	.031
B × C (interaction)	2	333561	3.78	.034
A × B × C (interaction)	2	247204	3.33	.05

A 2×2×3 ANOVA revealed a significant three-way interaction, $F(2, 30) = 3.33$, $p = .05$. To follow-up the three-way interaction, an Input Layout (2) × Taxi Clearance Length (3) ANOVA was conducted at each level of feedforward. Without feedforward, the Input Layout × Taxi Clearance Length interaction was not significant, $F(2, 30) = 1.76$, $p = .19$, however there was a main effect of input layout, $F(1,15) = 9.08$, $p = .009$, with softkeys yielding longer taxi clearance input times than QWERTY at all taxi clearance lengths (4-segment, $t(15) = 2.26$, $p = .039$, $d = 0.85$; 5-segment, $t(15) = 3.67$, $p = .002$, $d = 1.26$; 6-segment, $t(15) = 2.29$, $p = .037$, $d = 0.83$). In the feedforward conditions, however, the Input Layout × Taxi Clearance Length interaction was significant, $F(2, 30) = 6.10$, $p = .006$. Softkeys yielded a longer input time than QWERTY for the 6-segment taxi clearances, $t(15) = 2.37$, $p = .031$, $d = .71$, but not for the 4- or 5-segments taxi clearances, $t(15) = 0.71$, $p = .49$, $d = 0.13$ and $t(15) = 1.02$, $p = .33$, $d = 0.27$, respectively (see Figure 9).
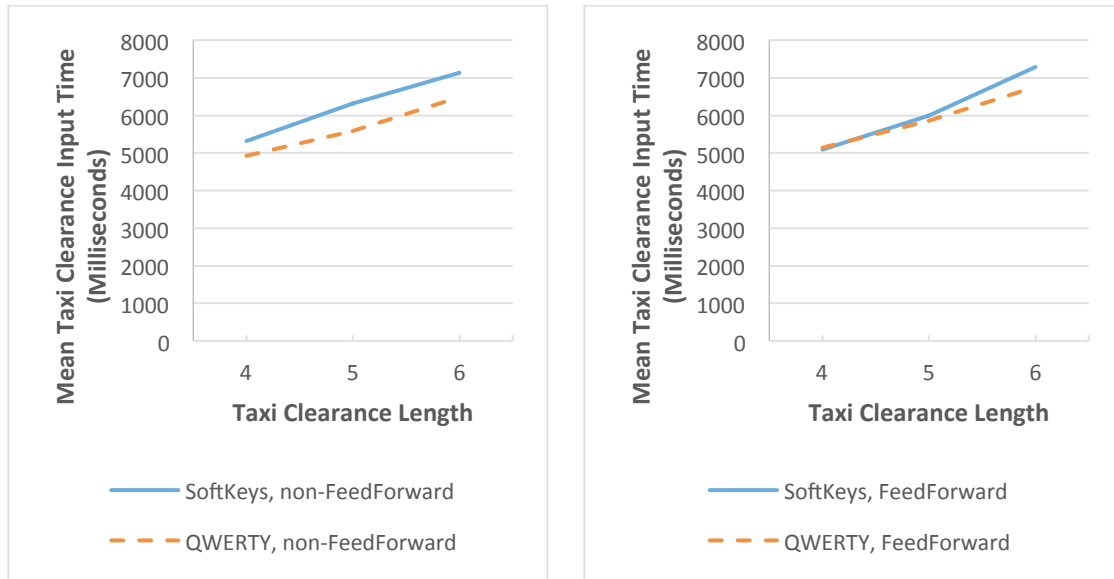
Figure 9. Interactions for taxi clearance input times: Input Layout (2) × Taxi Clearance Length (3).

The 2×2×3 interaction was also followed up with a Non-Feedforward/ Feedforward (2) × Taxi Clearance Length (3) ANOVA at each input layout (see Figure 10). In the softkeys layout, Non-Feedforward/Feedforward × Taxi Clearance Length interaction was significant, $F(2, 30) = 4.63$, $p = .018$. Feedforward had significantly shorter input times for 4-segment taxi clearances, $t(15) = 2.13$, $p = .05$, $d = 0.57$), but there was no significance at the 5- or 6-segment clearances, $t(15) = 1.86$, $p = .08$, $d = 0.58$ and $t(15) = 0.83$, $p = .42$, $d = 0.19$, respectively. In the QWERTY layout conditions, on the other hand, there was not a significant interaction for Non-Feedforward/Feedforward × Taxi Clearance Length, $F(2, 30) = 0.30$, $p = .74$.

Figure 10. Interactions for taxi clearance input times: Non-Feedforward/Feedforward (2) × Taxi Clearance Length (3).

To summarize, the QWERTY layout reduced taxi clearance input times relative to the softkeys layout (but when feedforward was present this effect was only observed for the longer, 6-segment, taxi clearances). As for the effect of feedforward, it was seen that with the softkeys input layout, feedforward reduced input times compared to non-feedforward, but only for the shorter four segment clearances. Feedforward did not impact input times for the QWERTY input layout.

Results for between-keystroke input times. Figure 11 summarizes the data collected for between-keystroke input times. Table 3 summarizes the results of the 2×2×3 (Input Layout × Non-Feedforward/Feedforward × Taxi Clearance Length) ANOVA.
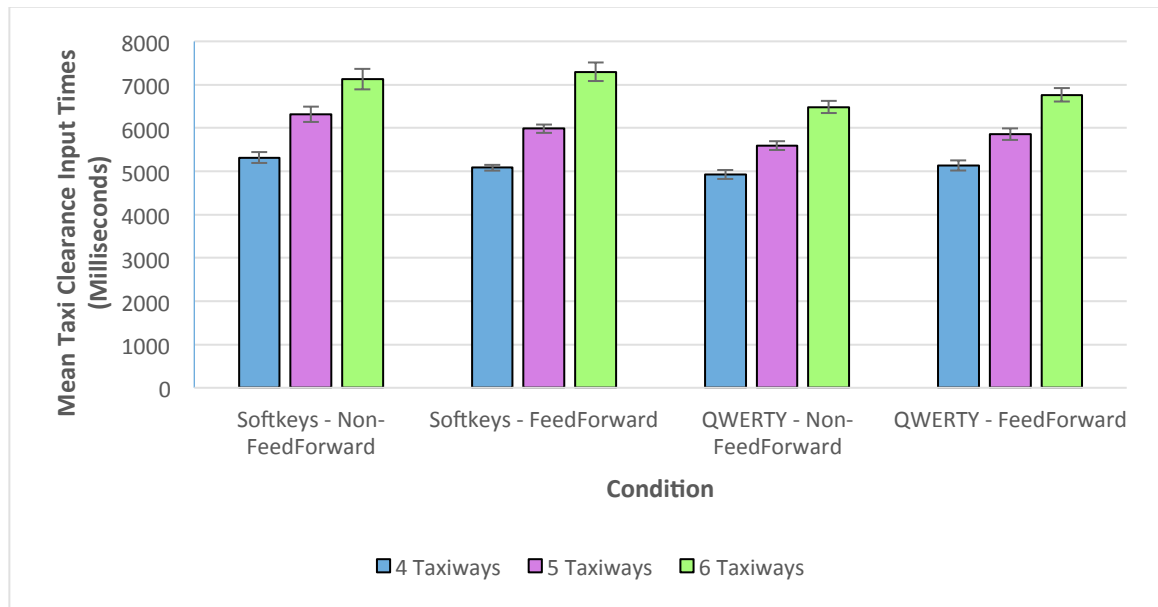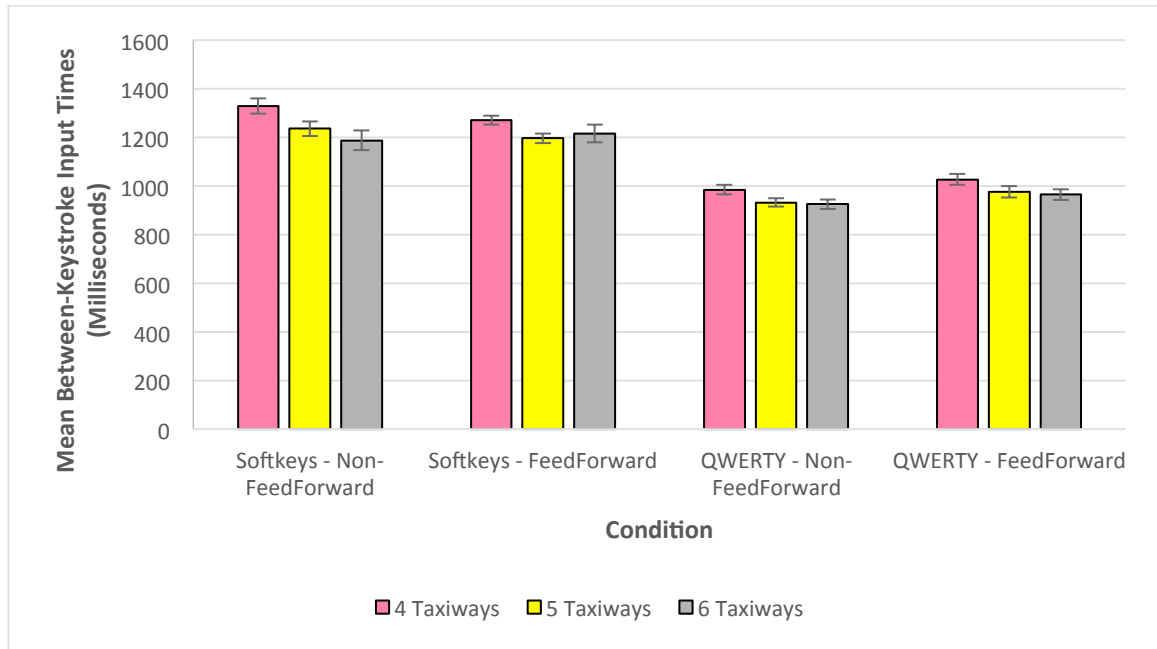
Figure 11. Mean between-keystroke input times. Shown with ±1 SE.

Table 3

Input Layout × Non-Feedforward/Feedforward × Taxi Clearance Length Repeated Measures ANOVA for Between-Keystroke Input Times

Source	df	MS	F	Sig.
(A) Input Layout	1	3528930	195.15	< .001
(B) Non-Feedforward/Feedforward	1	3970	0.49	.50
(C) Taxi Clearance Length	2	117959	26.52	< .001
A × B (interaction)	1	50574	2.27	.15
A × C (interaction)	2	6607	1.42	.26
B × C (interaction)	2	7628	3.71	.036
A × B × C (interaction)	2	9125	5.47	.009

A 2×2×3 ANOVA revealed a significant three-way interaction, $F(2, 30) = 5.47$, p = .009. The 2×2×3 interaction was followed up with a Non-Feedforward/Feedforward

(2) × Taxi Clearance Length (3) ANOVA at each input layout (see Figure 12). In the

softkeys layout, Non-Feedforward/Feedforward × Taxi Clearance Length interaction was

significant, $F(2, 30) = 6.01$, $p = .006$. Feedforward had significantly shorter between-

keystroke input times for 4-segment taxi clearances, $t(15) = 2.13$, $p = .05$, $d = 0.57$, but

there was no significance at longer taxi clearances (5-segment $t(15) = 1.23$, $p = .24$, $d =$

0.38; 6-segment $t(15) = 0.83$, $p = .42$, $d = 0.19$). In the QWERTY layout conditions,

there was not a significant interaction for Non-Feedforward/Feedforward × Taxi
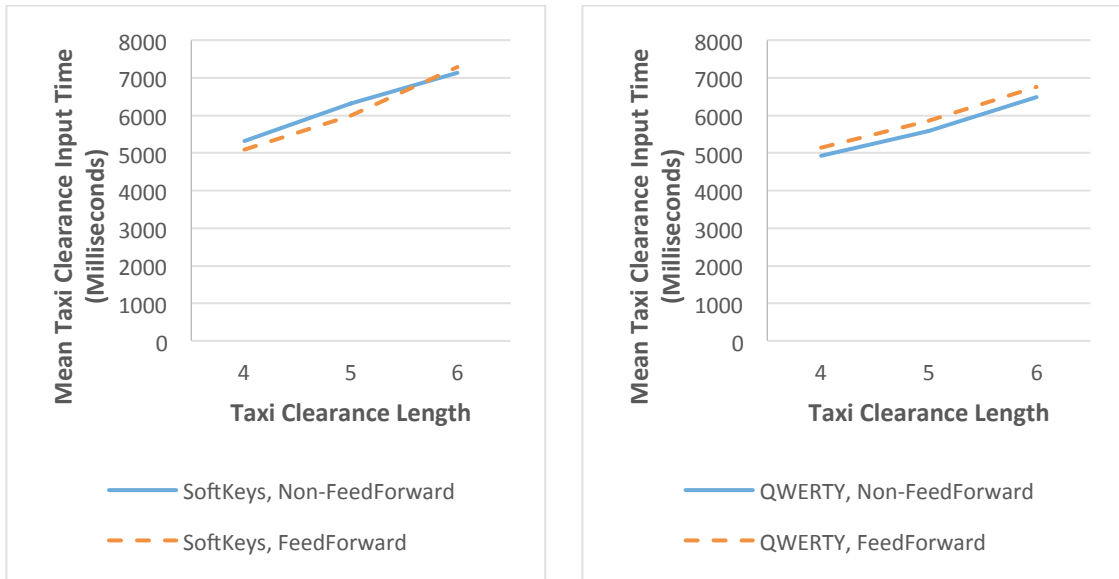
Clearance Length, $F(2, 30) = 0.06$, $p = .94$.
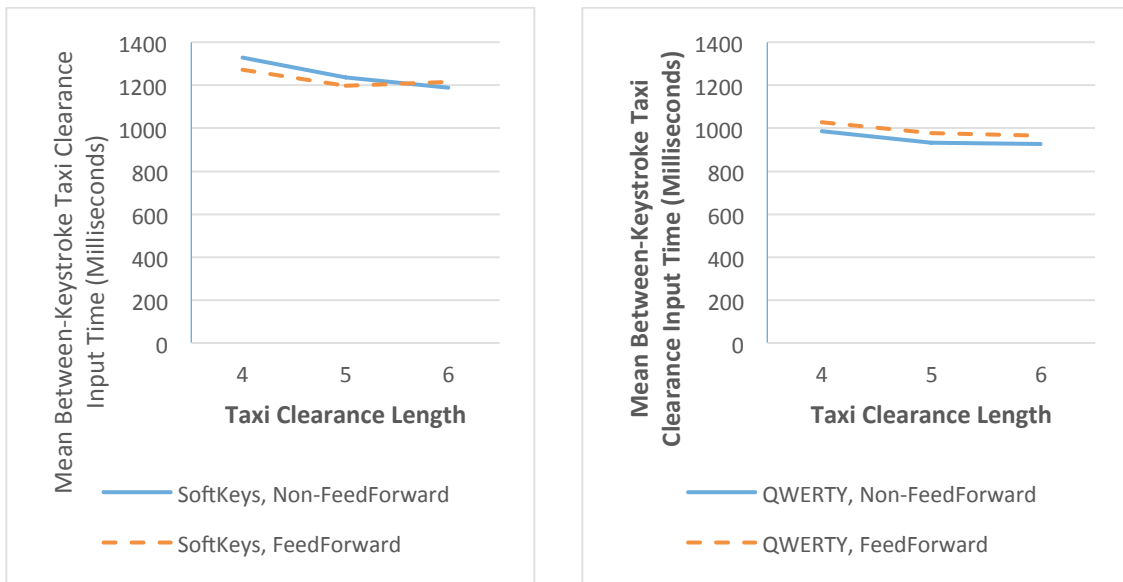
Figure 12. Interactions for between-keystroke input times: Non-Feedforward/
Feedforward (2) × Taxi Clearance Length (3).

Subjective Ratings

Participants were asked to provide various ratings addressing their perceived workload, ease of use and safety across conditions, as well as their ranking of conditions.

Results for perceived workload. After each experimental condition, participants rated seven workload dimensions (modified NASA-TLX) using a seven-point scale (1 = very low, 7 = very high). Results are shown in Figure 13. Individual 2×2 ANOVAs were conducted on each workload dimension and no significance was found across any of the workload dimensions (see Tables 4 through 10).
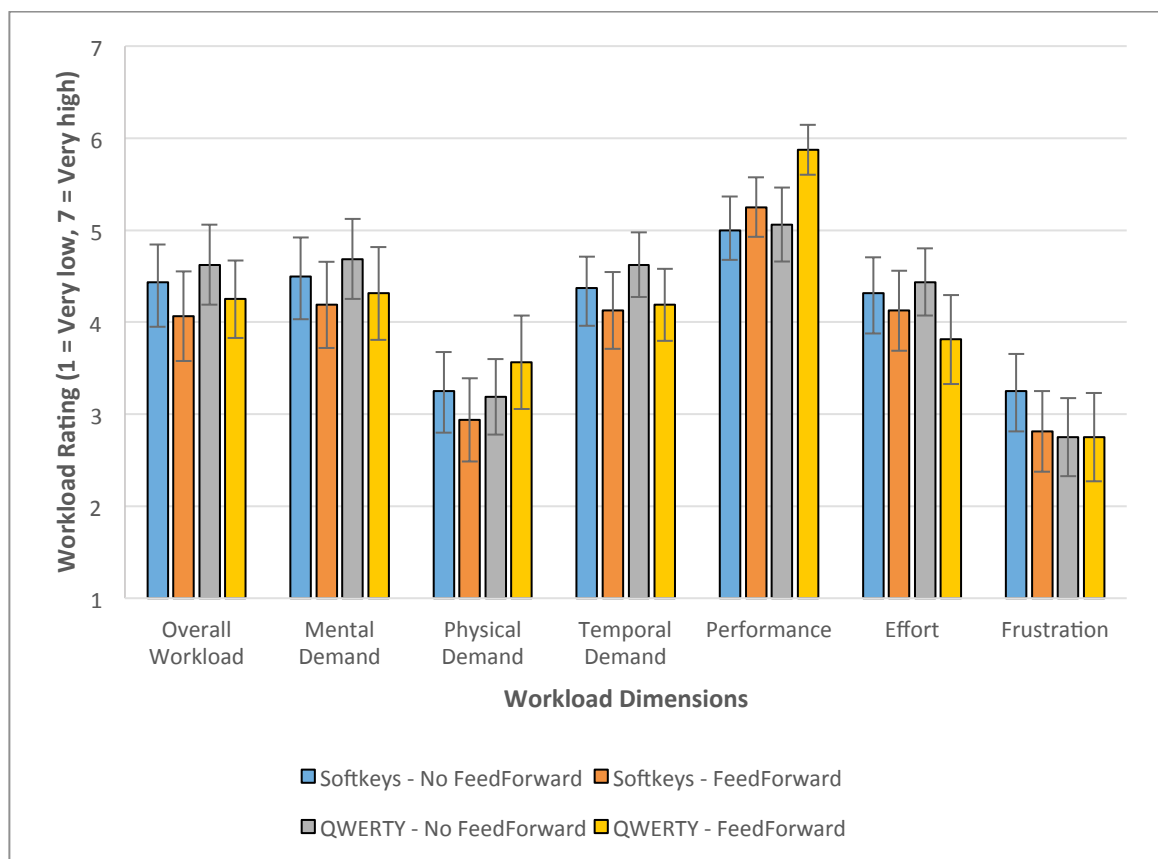
Figure 13. Mean ratings (±1 SE) from workload questionnaire.

Table 4

Input Layout × Non-Feedforward/Feedforward ANOVA for Overall Workload

Source	df	MS	F	Sig.
(A) Input Layout	1	0.56	0.28	.60
(B) Non-Feedforward/Feedforward	1	2.25	1.05	.32
A × B (interaction)	1	0.00	0.00	1

Table 5

Input Layout × Non-Feedforward/Feedforward ANOVA for Mental Demand

Source	df	MS	F	Sig.
(A) Input Layout	1	0.39	0.21	.66
(B) Non-Feedforward/Feedforward	1	1.89	0.77	.39
A × B (interaction)	1	0.02	0.01	.94

Table 6

Input Layout × Non-Feedforward/Feedforward ANOVA for Physical Demand

Source	df	MS	F	Sig.
(A) Input Layout	1	1.27	0.83	.38
(B) Non-Feedforward/Feedforward	1	0.02	0.01	.92
A × B (interaction)	1	1.90	1.55	.23

Table 7

Input Layout × Non-Feedforward/Feedforward ANOVA for Temporal Demand

Source	df	MS	F	Sig.
(A) Input Layout	1	0.39	0.25	.63
(B) Non-Feedforward/Feedforward	1	1.89	2.12	.17
A × B (interaction)	1	0.14	0.08	.78

Table 8

Input Layout × Non-Feedforward/Feedforward ANOVA for Performance

Source	df	MS	F	Sig.
(A) Input Layout	1	1.89	1.79	0.20
(B) Non-Feedforward/Feedforward	1	4.52	3.35	0.09
A × B (interaction)	1	1.27	1.23	0.29

Table 9

Input Layout × Non-Feedforward/Feedforward ANOVA for Effort

Source	df	MS	F	Sig.
(A) Input Layout	1	0.14	0.12	.73
(B) Non-Feedforward/Feedforward	1	2.64	1.97	.18
A × B (interaction)	1	0.77	0.77	0.40

Table 10

Input Layout × Non-Feedforward/Feedforward ANOVA for Frustration

Source	df	MS	F	Sig.
(A) Input Layout	1	1.27	0.76	.40
(B) Non-Feedforward/Feedforward	1	0.77	0.39	.54
A × B (interaction)	1	0.77	0.26	.62

Results for condition rankings. In the post-study questionnaire, participants were asked to rank the five conditions in order of preference (this included the written condition, which is not discussed in other analyses of this paper, but necessary for comparing overall condition ranks). The most top-preferred condition was QWERTY with feedforward. The softkeys condition with non-feedforward was the only condition not to be ranked first by anyone. Figure 14 summarizes these findings.
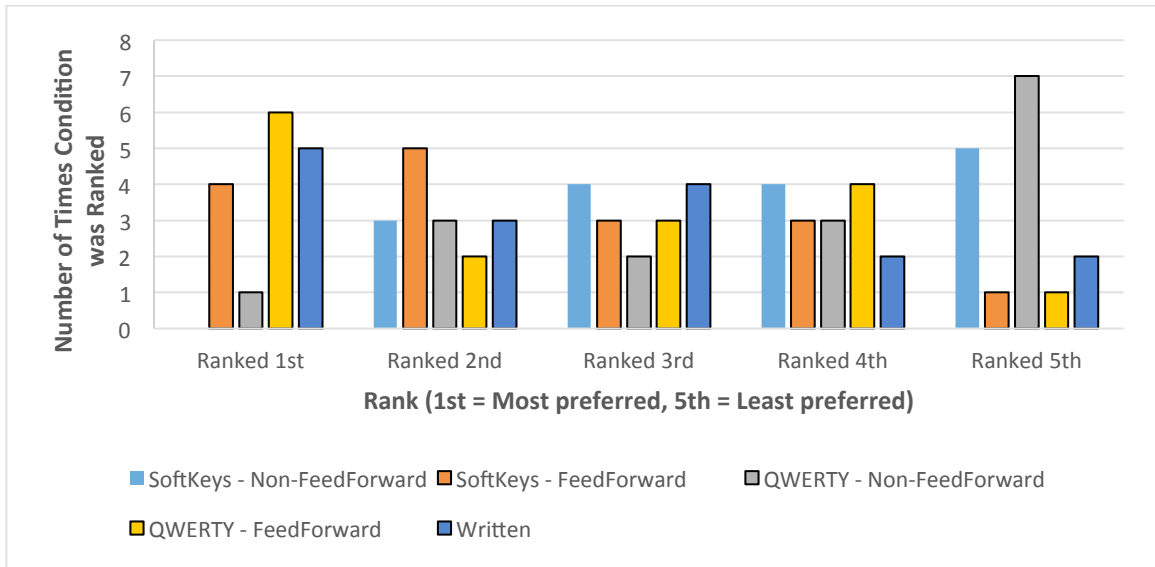
Figure 14. Number of times each condition was ranked 1st, 2nd, 3rd, 4th, and 5th.

Results for ease of use. In the post-study questionnaire, participants were asked to rate the ease of use of each of the conditions from 1 to 5 (where 1 = Very difficult to use, 5 = Very easy to use). Mean ratings for ease of use are presented in Figure 15. A 2×2 ANOVA (see Table 11) revealed a significant interaction, $F(1, 15) = 4.36$, $p = .05$. Post-hoc analyses revealed no difference between non-feedforward and feedforward in the softkeys layout, $t(15) = 1.78$, $p = .10$. However, the non-feedforward QWERTY condition yielded significantly lower ratings for ease of use than did the feedforward QWERTY condition, $t(15) = 3.126$, $p = .006$, $d = 0.87$. Post-hoc analyses also showed that there were no significant differences in ratings between either the softkeys and QWERTY layouts without feedforward or the softkeys and QWERTY layouts with feedforward. These findings suggest that feedforward improved ease of use for the QWERTY method, but did not support improved ease of use for the softkeys method.
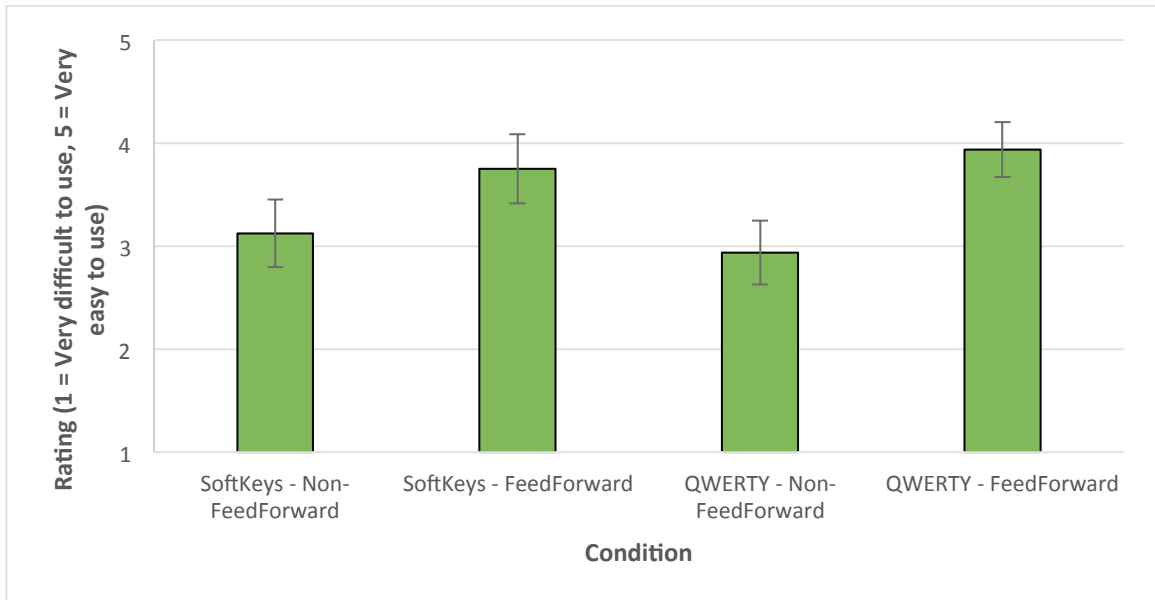
Figure 15. Mean ratings (±1 SE) for ease of use across the four conditions.

Table 11

Input Layout × Non-Feedforward/Feedforward ANOVA for Ease of Use

Source	df	MS	F	Sig.
(A) Input Layout	1	0.00	0.00	1
(B) Non-Feedforward/Feedforward	1	10.56	6.35	.02
A × B (interaction)	1	0.56	4.36	.05

Results for safety. In the post-study questionnaire, participants were asked to rate safety (if they had to input a taxi clearance while taxiing) across the four conditions. Note, participants were asked to rate (on a scale of 1 to 5) the safety "while taxiing" in the event of a route amendment during taxi (where 1 = Very unsafe and 5 = Very safe). For the most part, however, proper pilot procedure would dictate that the pilot taxiing would stop before entering the taxi clearance in a head-down device. Mean ratings for safety are presented in Figure 16. A 2×2 ANOVA (see Table 12) revealed that there was no

significant interaction, $F(1, 15) = .32$, $p = .58$. The ANOVA also showed that there was

not a significant main effect of layout (softkeys vs. QWERTY), but there was a

significant main effect of non-feedforward/feedforward, with non-feedforward receiving

lower ratings for safety, $F(1, 15) = 15.63$, $p = .001$. These findings suggest that

participants found the feedforward conditions to be safer than non-feedforward

conditions.
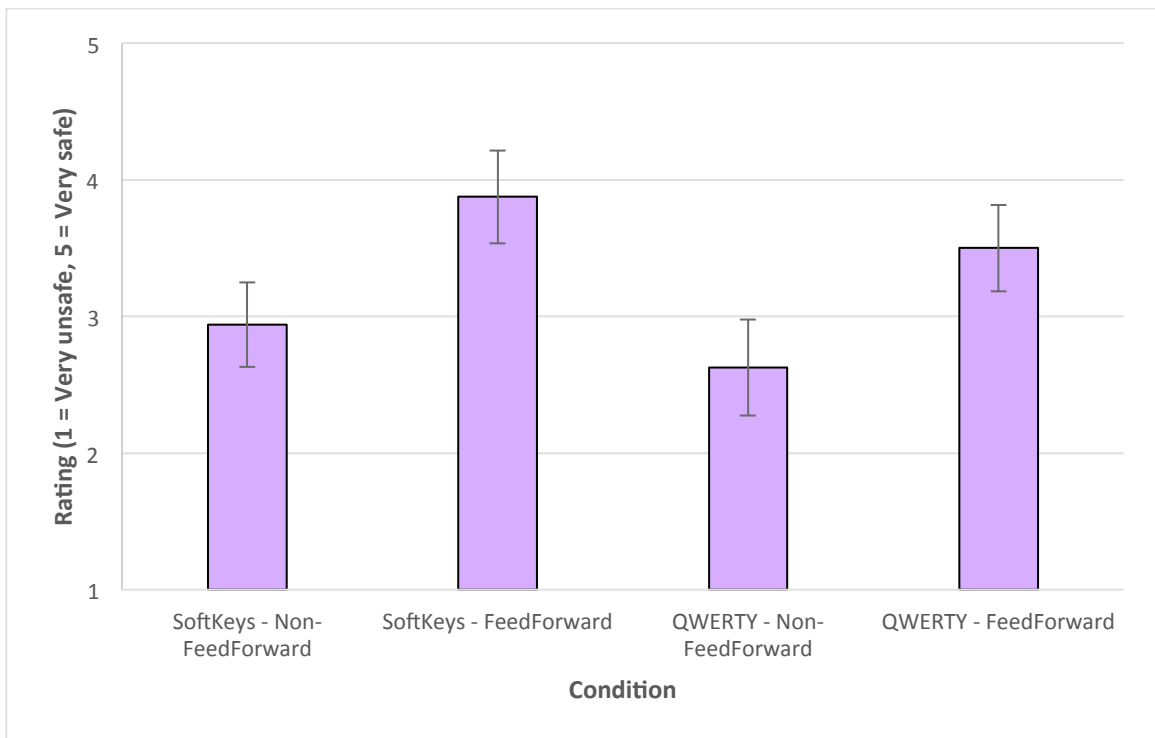
Figure 16. Mean ratings (± 1 SE) for safety, if participants had to input a taxi clearance
while taxiing.

Table 12

Input Layout × Non-Feedforward/Feedforward ANOVA for Safety

Source	df	MS	F	Sig.
(A) Input Layout	1	1.89	1.68	.21
(B) Non-Feedforward/Feedforward	1	13.14	15.63	.001
A × B (interaction)	1	0.02	0.32	.58

Discussion

It has been suggested that displaying the cleared taxi route to pilots on an AMM can enable conformance monitoring and reduce taxi navigation errors (Hooey & Foyle, 2006; Theunissen et al., 2007), while also reducing head-down time and workload (Schönefeld & Möller, 2012). In 2007, Theunissen et al. emphasized that taxi route input methods should allow pilots to input taxi routes accurately, quickly, and with workload low enough such that they can complete the task as the clearance is being read by ATC. Such methods would also allow for pilots to read clearances back to ATC, which closes the integrity loop and allows pilots and ATC to cross-check. A similar version of the virtual keys that Theunissen et al. (2007) tested was used in the present study, as well as a QWERTY layout (intended to alleviate the potential issue of limited keys at complex airports).

The results of the present study supported the hypothesis that QWERTY would yield faster input times than softkeys, but there was only significance for longer (6-segment) taxi clearances with feedforward; however, the study did not support the hypothesis that QWERTY would yield a lower frequency of input errors than softkeys.

In fact, for shorter (4-segment) taxi clearances with feedforward, the results negated the hypothesis that QWERTY would produce fewer input errors than softkeys. It is possible that this may have to do with the fact that there was much more horizontal space between the keys on the softkeys layout than on the QWERTY layout.

This study also tested a context-sensitive input function (called feedforward) in which only valid taxiways connected to the last entered leg are available for the user to select, similar to that tested by Theunissen et al. (2007). The present study did not support the hypotheses that input times and Error Rate would both be lower for feedforward than non-feedforward conditions. The feedforward feature was meant to serve as an aide for the participants, but a number of participants commented that the contrast between the gray and black keys may have been too great. They noted that the constant switching of black and gray keys made the layouts look chaotic. It is especially odd that given the handful of participants that made this claim, QWERTY with feedforward was ranked as the top preference for input layout the most. Future research should look into contrasting colors and how strong the difference between keys that are "viable/not viable" should be.

The ratings for safety seemed rather low in this study, and that may be due to poor wording of the question in the post-study questionnaire, which asked users to rate the safety of each condition, if they had to enter taxi clearances while taxiing. It should have been clarified that hypothetically, the task of inputting taxi clearances would only be done while the aircraft was stopped (at an intersection, for example). It is possible that participants assumed the rating referred to taxiing as the airplane being in motion, which

35

could potentially be interpreted as a situation similar to texting while driving—an action that is widely understood to be unsafe.

Speed, accuracy and workload are important when it comes to considering taxi clearance input layouts (Theunissen et al., 2007). workload ratings showed no interactions between input layout or non-feedforward/feedforward, and there were no concerning workload ratings. Performance yielded fairly high ratings, which suggests that participants felt that they did well. Perhaps larger keys and/or spacing between keys can be tested in future studies to determine whether the high proportions of errors when using the QWERTY layout are due to spacing and/or key size. In addition, the present study neither allowed participants to delete characters one by one (participants were forced to clear an entire string if they noticed an input mistake and needed to retype it correctly) nor gave them the ability to ask for a taxi clearance to be repeated, due to software limitations. Such limitations are not realistic, and in fact would very likely increase user workload since entire strings have to be recalled and memorized if an input error is made.

Future research may also be directed toward comparing other alphanumeric input layouts in addition to (or aside from) the ones used in the present study (such as alphabetically ordered keys or "next page" options), since only two input layouts were compared. In addition, creating an overall more realistic flight-deck environment for participants would be ideal (e.g., more realistic environment, or at least devices that are more realistic to pilots inputting data on flight-deck interfaces). Due to resource limitations, the present study utilized a 21.5" touchscreen monitor, which utilized only a

small area of the screen. In terms of creating a more realistic environment, future studies

might be conducted in flight-deck simulators so that the results can better justify

implementing certain input layouts on flight decks. The present study was also not

realistic in terms of the number of trials and the three-second pause between trials (since

it would be ideal for pilots to input a taxi clearance just once). This may have had an

effect on the workload ratings, since users would not generally expect a new taxi

clearance every three seconds. Future studies might incorporate taxi route input as just

one element in a series of tasks for surface operations.

The results of the present study may be considered in advancing future research

that focuses on providing taxiing and airport information during surface operations. On a

broader scale, the results of this study have also contributed the domains of input layouts,

flight-deck operations, and surface operations. As we see more and more people using

touchscreen devices such as tablets, desktops, laptops and phones, QWERTY has become

the norm for inputting alphanumeric characters, and this study reflects that notion. Flight

deck displays are no exception for "frequently-used displays" amongst pilots, so it could

benefit them to have interfaces with touchscreen (for purposes of easy re-programming)

QWERTY keyboard. Based on the results and analyses of the present study, the author

suggests that for taxi clearance Inputs, (1) QWERTY remain the standard for

alphanumeric inputs and (2) feedforward be investigated further, with a focus on

participant preference and performance of black-gray contrast of keys.

References

Ahlstrom, V., & Longo, K. (2003). Chapter 9: Input Devices. In *Human Factors Design Standard* (HF-STD-001) (9-4–9-6). Atlantic City International Airport, NJ: Federal Aviation Administration William J. Hughes Technical Center.

American National Standards Institute [ANSI] (1988). American national standard for human factors engineering of visual display terminal workstations (ANSI/HFSStandard No. 100). Santa Monica, CA: The Human Factors Society.

Department of Energy (1992). Human factors engineering design criteria (DOE-HFAC1). Washington, DC: United States Department of Energy.

Ha, V., Inkpen, K. M., Mandryk, R. L., & Whalen, T. (2006). Direct intentions:

The effects of input devices on collaboration around a tabletop display. Proceedings from IEEE: *First IEEE International Workshop on Horizontal Interactive Human-Computer Systems*, 177-184.

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.), *Human Mental workload.* Amsterdam: North Holland Press.

Hooey, B.L., & Foyle, D.C. (2006). Pilot navigation errors on the airport surface: Identifying contributing factors and mitigating solutions. *The International Journal of Aviation Psychology, 16*(1), 51-76.

Howard, B. (2014). Ford admits touchscreen defeat, puts the buttons and knobs back into Ford Sync. Retrieved from http://www.extremetech.com/extreme/176936-ford-admits-touchscreen-defeat-puts-the-buttons-and-knobs-back-into-ford-sync

Joint Planning and Development Office (2011). Overview of NextGen concepts and capabilities. Concept of Operations for the Next Generation Air Transportation System, 3.2, 1-10.

Kiljander, H. (2004). Evolution and usability of mobile phone interaction styles (Doctoral dissertation). Retrieved from Aalto University Library. (ISBN 951-22-7320-9)

Matthews, L.L., Bryant, D.J., Webb, R.D.G., & Harbluk, J.L. (2001). Model for Situation Awareness and Driving: Application to Analysis and Research for Intelligent Transportation Systems. *Transportation Research Record, 1779*, 26-32.

Midkif, A. H., Hansman, R. J., & Reynolds, T. G. (2004). Air carrier flight operations (ICAT-2004-3). Cambridge, MA: MIT International Center for Air Transportation.

National Institute of Standards and Technology (2013). *Set your computer clock via the Internet using tools built into the operating system.* Retrieved from http://www.nist.gov/pml/div688/grp40/its.cfm

Nielsen, J. (1995). Advances in human-computer interaction (vol. 5). Norwood, NJ: Ablex Publishing.

Norman, D.A. (2002). *The design of everyday things.* New York, NY: Basic Books.

Regan, M.A., Lee, J.E., & Young, K.L. (Eds.) (2009). *Driver distraction: Theory, effect and mitigation.* Boca Raton, FL: CRC Press.

Roefs, F., Theunissen, E., & Koeners, J. (2007). Manual versus speech input for entering a taxiroute into an electronic flight bag. Proceedings from HFES: *Human Factors and Ergonomics Society Annual Meeting, 51*(2), 60-64.

Schönefeld, J., & Möller, D. P. F. (2012). Runway incursion prevention systems: A review of runway incursion avoidance and alerting system approaches. *Progress in Aerospace Sciences*, *51*, 31-49.

Theunissen, E., Roefs, F., Koeners, G. J. M., & Bleeker, O. F. (2007). Design, integration and evaluation of an application for input of taxiroutes. Proceedings from Digital Avionics Systems Conference: *IEEE/AIAA 26th*, 6.B.1-1 - 6.B.1-11.

Theunissen, E., Roefs, F. D., Koeners, G. J. M., & Bleeker, O. F. (2008). Taxi route input-specification or selection? Proceedings from Digital Avionics Systems Conference: *IEEE/AIAA 27th*, 4.A.3-1 - 4.A.3-10.

Theunissen, E., Roefs, F. D., Moncur, B., & Jinkins, R. D. (2006). Error detection and correction for speech recognition using airport layout information: Concept and evaluation. Proceedings from Digital Avionics Systems Conference: *IEEE/AIAA 25th*, 1-11.

Truitt, T.R., & Muldoon, R. (2010). Data communications segment 2 airport traffic *control tower human-in-the-loop simulation* (DOT/FAA/TC-10/05). Atlantic City International Airport, NJ: Department of Transportation, Federal Aviation Administration.

Appendix A: NASA Ames Human Research Institutional Review Board Letter of Approval

Subject: HRIRB Approval of HRII-13-13

Date: Wednesday, June 12, 2013 3:36:20 PM PT

From: Timm, Elaine M. (ARC-D)[Lockheed Martin Space OPNS]

To: Foyle, David C. (ARC-TH)

CC: Pelligra, Ralph (ARC-Q), Hooey, Becky L. (ARC-TH)[SAN JOSE STATE UNIVERSITY FOUNDATION INC]

The following message is sent on behalf of Ralph Pelligra, M.D., Chair, Human Research Institutional Review Board (HRIRB):

Dear Dr. Foyle:

Protocol **HRII-13-13**, "Flight/Taxi Displays," (formerly HRII-12-15), was approved for renewal by the NASA Ames Human Research Institutional Review Board (HRIRB) at the monthly meeting on June 06, 2013. You are authorized to conduct research activities subject to Ames Procedural Requirements (APR) 7170.1, Human Research Planning and Approval. Your approval period is valid from **June 07, 2013** through **June 06, 2014**. Please note the termination date on your calendar. Any modifications to an approved protocol require further approval by the Chair, HRIRB and a designated HRIRB member.

Please retain a copy of this confirmation message for your files. Begin a renewal request six weeks prior to **June 06, 2014**, the protocol's expiration date. Allow two to three weeks processing in the Office for the Protection of Research Participants (OPRP), prior to an HRIRB meeting. If you do not renew the protocol, submit the end of study report within 30 days of the protocol's expiration. Both forms are available on the HRIRB web site, http://hrirb.arc.nasa.gov.

If you have questions, contact me at 650.604.5163, or by e-mail, Ralph.Pelligra-1@nasa.gov. Immediately notify me of any injury to a human participant, whether or not it was considered a risk inherent to the research. Injury includes, but is not limited to, bodily harm, psychological trauma or release of potentially damaging personal information.

Good luck in your research efforts.

Ralph Pelligra, M.D.
Chair, HRIRB
By
Elaine Timm
Recorder, Human Research Institutional Review Board;
Office for the Protection of Research Participants
M/S 243-2
Bldg. 243, Room 120
voice: 650.604.0119
fax: 650.604.6233
E-mail: Elaine.M.Timm@nasa.gov

Appendix B: San José State University Human Subjects Institutional Review Board Letter of Approval

SAN JOSÉ STATE
UNIVERSITY

Division of Academic Affairs

Associate Vice President
Graduate Studies & Research

www.sjsu.edu/gradstudies

One Washington Square
San José, California 95192-0025
Voice: 408-924-2427

www.sjsu.edu

To: Lara Cheng

From: Pamela Stacks, Ph.D.
        Associate Vice President
        Graduate Studies and Research

Date: June 17, 2014

The Human Subjects-Institutional Review Board has approved your request to use human subjects in the study entitled:

"Comparing Input Methods of Taxi Routes on Flight Deck Control Display Units"

This approval is contingent upon the subjects participating in your research project being appropriately protected from risk. This includes the protection of the confidentiality of the subjects' identity when they participate in your research project, and with regard to all data that may be collected from the subjects. The approval includes continued monitoring of your research by the Board to assure that the subjects are being adequately and properly protected from such risks. If at any time a subject becomes injured or complains of injury, you must notify Dr. Pamela Stacks, Ph.D. immediately. Injury includes but is not limited to bodily harm, psychological trauma, and release of potentially damaging personal information. This approval for the human subject's portion of your project is in effect for one year, and data collection beyond June 17, 2015 requires an extension request.

Please also be advised that all subjects need to be fully informed and aware that their participation in your research project is voluntary, and that he or she may withdraw from the project at any time. Further, a subject's participation, refusal to participate, or withdrawal will not affect any services that the subject is receiving or will receive at the institution in which the research is being conducted.

If you have any questions, please contact me at (408) 924-2427.

Protocol #S1402099

cc. Sean Laraway   0120

Appendix C: Short Description of Study and Consent Form Provided to Participants

Purpose of this Study

This study will attempt to unveil insight on whether our proposed designs for modern software interfaces can increase pilot performance of Control Display Unit (CDU) use (in particular for inputting taxi routes) with little to no hindrance on safety, workload, or situation awareness.

Appendix D: Workload Questionnaire

Workload

Overall Workload: How would you rate the overall workload in this condition? *

	1	2	3	4	5	6	7	
Very Low	○	○	○	○	○	○	○	Very High

Mental Demand: How mentally demanding was the task? *

	1	2	3	4	5	6	7	
Very Low	○	○	○	○	○	○	○	Very High

Physical Demand: How physically demanding was the task? *

	1	2	3	4	5	6	7	
Very Low	○	○	○	○	○	○	○	Very High

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

	1	2	3	4	5	6	7	
Very Low	○	○	○	○	○	○	○	Very High

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

	1	2	3	4	5	6	7	
Very Low	○	○	○	○	○	○	○	Very High

Frustration: How insecure, discouraged, irritated, stressed, and annoyed were you? *

	1	2	3	4	5	6	7	
Very Low	○	○	○	○	○	○	○	Very High

Submit

Never submit passwords through Google Forms.

Appendix E: Description of Selection Process for Software used in Study

Selecting a program to build and run the study on

When selecting program to use for experiment-building and data collection there were a few desired criteria: no programming language or scripting necessary, can log user inputs to the millisecond, and is compatible with touchscreens. The desire for "no programming language or scripting necessary simply stemmed from the researcher's lack of experience in the area. Thus, it would have been ideal to build experiments with primarily "drag-and-drop" functions. Millisecond input logging was also a very important feature for this experiment, since participants were to be inputting within rather small, constrained areas and large differences in input times were not expected. Finally, it was important that the program be compatible with touchscreen, since the present study pushed for modernizing flight-deck hardware (i.e., moving away from hardware keys and buttons to touchscreen/modifiable keys and buttons).

While evaluating various programs that would be able to fulfill the requirements described above, options for combining various programs as well as options for using a single program were considered. It was determined early on in this process that there are very few options, at this time, for script-free experiment-building tools combined (or combinable) with data input collection. The program that seemed to be able to fulfill this study's requirements was Paradigm, a simple experiment-building, data-collecting software. The option to write script for Paradigm-built experiments exists, but the Paradigm website advertises the product as being a no-script-necessary experiment-building tool.

In addition, the initial proposal for this study called for a touchscreen tablet (such as a Samsung Galaxy 10.1 or an Apple iPad Air), a capability that Paradigm was advertised as being capable of. However, this became an issue when the researcher learned through experiment-building and Customer Support that the functions needed for the study require script. On top of that, Customer Support revealed to the researcher that script is not compatible with mobile. The solution was to use a touchscreen monitor, connected to a laptop (MacBook Air).

A second laptop was needed for the experiment (Asus VivoBook) because the taxi clearance audio clips would stop playing once the user started typing (the interface built using Paradigm could not be dynamic, so each input that the participant made was essentially a "page flip"). The purpose of playing the taxi clearance audio clips on a second laptop using Paradigm was to collect the timestamps for when each trial essentially started for the participant. This would allow the researcher to determine how long it took for participants to begin inputting the taxi clearance from the time the audio started.

Demographics Questionnaire

* Required

Gender *

○ Male

○ Female

Age *

_____

Do you currently have a valid GA license? *

○ Yes

○ No

When as the last time you flew an aircraft? *

mm/dd/yyyy

How many total flying hours have you logged? *

_____

What other types of input devices have you used on the flight deck (tablets, smart phones, stand-alone GPS units, etc.)? *

What, if any, navigation aids do you use during taxi? What is the input method for each of these devices? Please list. *

Referring to the previous question, how often do you use these devices (e.g., everyday, once a month)? *

On average, how many flights do (or did) you make in one month? *

Do you own a touchscreen tablet or touchscreen computer? If so, what model(s) and size(s) (e.g., iPad Mini, Samsung Galaxy 10.1)? Please list. *

If applicable, when was the last time you used your touchscreen device(s)? *

mm/dd/yyyy

If applicable, how often do you use a tablet or touchscreen computer (e.g., everyday, a few times a week, once per month)?

What navigation apps, if any, do you use for driving? How often do you use it/them (e.g., everyday, a few times a week, once per month)? Please list.

e.g., Google Navigation - twice a week

Submit

Never submit passwords through Google Forms.

Appendix G: Post-Study Questionnaire

Participant #: _____

Post-Study Questionnaire

1. Please rank the 5 input methods, 1 through 5, that you used today in order of overall personal preference (1 = most preferred, 5 = least preferred).

SoftKeys without FeedForward
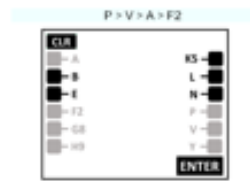
RANK: _____

SoftKeys with FeedForward
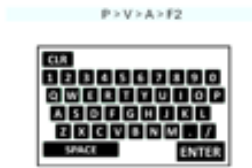
RANK: _____

QWERTY without FeedForward

RANK: _____

QWERTY with FeedForward
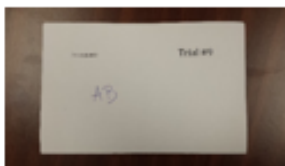
RANK: _____

Written

RANK: _____

2. Please rate the ease of use of each of the conditions.

	Very difficult to use		Neither easy nor difficult to use		Very easy to use
SoftKeys without FeedForward	1	2	3	4	5
SoftKeys with FeedForward	1	2	3	4	5
QWERTY without FeedForward	1	2	3	4	5
QWERTY with FeedForward	1	2	3	4	5
Written	1	2	3	4	5

3. Please rate the safety of each condition, *if you had to input a taxi clearance while taxiing*.

	Very unsafe		Neither safe nor unsafe		Very Safe
SoftKeys without FeedForward	1	2	3	4	5
SoftKeys with FeedForward	1	2	3	4	5
QWERTY without FeedForward	1	2	3	4	5
QWERTY with FeedForward	1	2	3	4	5
Written	1	2	3	4	5

4. Please compare the four touchscreen conditions to the Current Day (written) condition.

	Was it easier or harder to use than writing?		Was it faster or slower than writing?		Was it prone to more or less error?	
	Easier	Harder	Faster	Slower	More	Less
QWERTY without FeedForward						
QWERTY with FeedForward						
SoftKeys without Feedforward						
SoftKeys with FeedForward						

5. Do you have any comments or concerns about the study? If so, please describe.

_____

_____

_____

_____

_____

_____

_____

_____

_____

Appendix H: Researcher Script

SCRIPT

- Introduction
    - Thank you for taking the time to come in and participate in this study, I really appreciate it.
    - Here's the consent form—please read through it and sign it, and let me know if you have any questions.
    - In this experiment I'm comparing a few different interfaces for entering taxi clearances.
    - If in the future you could enter the taxi clearance into the FMS, for example, maybe you'd have on-board nav guidance for taxiing in airports (which might be helpful if you're not familiar with an airport's layout.
    - Are some interfaces going to be more reliable, easier to use, faster to use, than others? Those are the things I'm looking at.
    - The interfaces you'll be seeing today are simple, but a push toward modernizing the cockpit, really bringing it up to date.
    - Today you'll be entering taxi clearances, using five different interfaces. Each interface is its own condition, and each condition will have a set number of trials.
    - At the start of each trial, you'll see the display that you'll be doing the input on, and you'll hear your taxi clearance. Here's what the audio for the taxi clearances will sound like today. *Play audio.*
    - When you hear the taxi clearance, you're going to enter it into the touchscreen— you can enter it while it's being played. You'll see your input typed out on the screen.
    - One thing to note is that each letter is its own segment of the taxiway, except for letters followed by a number. These alpha + number pairs are also one segment. You'll see that when you're inputting the taxi route.
    - Do NOT type the "taxi via." You also won't need to worry about call signs or entering runways today.
    - When you're done entering it, you'll press "enter" and there will be a blank white screen with a three-second pause.
    - After the three seconds, the display will come up again and you will hear the next taxi clearance audio.
    - If you notice a mistake in your taxi clearance input, press the "clear" button. That will clear the input you've made, and you will re-enter the taxi clearance to the best of your ability and press "enter" when complete.
    - I'll show you a couple trials as you start each new condition, and you'll have four practice trials. If you have any questions after the practice trials and before starting the actual trials, I'll be able to answer them.
    - Between each condition (or set of trials), you'll have a five-minute break to get up, stretch, use the restroom, etc.
    - The goal is to enter or write down (depending on the condition) the ATC taxi clearance as **quickly and accurately** as possible, since the safety and efficiency of airport taxiing depends on those two things. However, since accuracy is more important than speed, you should prioritize accuracy. Do you have any questions so far?

- Touchscreen Conditions
  - *Open correct condition without entering any participant info.*
  - Here's the (first, second, etc.) condition you'll be doing. I'll go ahead and briefly demonstrate how to use it. While I'm demonstrating, feel free to stop me if you have any questions.
  - I'm going play an audio clip of a taxi clearance. It will sound just like a typical taxi clearance from a ground controller, only today we won't be using call signs. You'll type it in, and press "enter" when you're done.
  - **If *feedforward* condition:** In this condition you'll see gray keys/buttons and black keys/buttons. The black ones are "valid" inputs, meaning they intersect with the last taxiway you entered and you're able to continue the route. The gray ones are "invalid" inputs, meaning it's not possible to directly connect the last taxiway you entered with them. So you'll only be using the black keys.
  - **If *softkeys:*** The order of these keys is alphabetical, and the letters will never shift around on the display.
  - Now, I am going to demonstrate what a trial will look like. *Play audio on VivoBook, enter route correctly.*
  - Now I've input the taxi clearance. Notice this alpha + number pair. As I explained before, that's one segment and all these other individual letters are their own segment as well.
  - Now I'm going to press "enter," and here is the three-second pause before the next trial starts.
  - Now the display is back and the next trial will begin, and the next taxi clearance audio will play. This time I'm going to make a mistake and press "clear." Now I'll re-enter the route correctly, to the best of my ability, and press "enter" and wait for the next trial.
  - Do you have any questions?
  - Now you'll do four practice trials.
  - **For the second one,** before hitting "enter," go ahead and "clear" the route and re-enter it—just to get the hang of it.
  - Do you have any questions before we begin the actual trials?
- Current-Day Condition
  - *Open current-day condition without entering any participant info.*
  - This is the current-day condition. At the start of each trial, this touchscreen will only have an "enter" button on it, and I will play the taxi clearance audio.
  - You'll use the scratchpad provided to write down the clearance as you hear it, just as you might in current-day operations.
  - If you need to, you can make any corrections on the scratchpad
  - I'm going play an audio clip of a taxi clearance, and you'll write it down just like you would in current-day situations.
  - When you are done writing the taxi clearance, press "enter" and flip to the next page on your scratchpad.
  - There will be a three-second blank screen before the next trial begins.
  - Do you have any questions?
  - We'll do four practice trials now.
  - *After practice trials, ask if they have any questions.*

55

- Post-study
  - *Pull up post-study questionnaire and demographics questionnaire.*
  - You've completed all five of the conditions now, here is a post-study questionnaire I'd like you to complete, then a demographics questionnaire.
  - Here's the post-study questionnaire. If you need clarification on any of it, just let me know.
  - Here's the demographics questionnaire. *Open for participant.* Again, if you have any questions, just let me know. *Participant completes questionnaire, and then I check in the live form via MacBook that each question has a response.*
  - Thank you so much for coming in today, I really appreciate your time and participation.

Appendix I: Researcher Checklist

Experimenter's Checklist

Each morning

____ Screen resolution (Participant Touchscreen)

____ Monitor on tape marks

____ JAMBOX volume

____ JAMBOX on tape marks

____ JAMBOX plugged into MacBook

____ Vivobook 📋 (Ctrl) ▾

____ MacBook plugged into wall

____ VivoBook plugged into wall

____ Consent form

____ Scratchpad

Before each block

	Block 1	Block 2	Block 3	Block 4	Block 5
Audio condition open					
Input condition open					
Open post-study questionnaire	x	x	x	x	
Open demographics questionnaire	x	x	x	x	

After each block

	Block 1	Block 2	Block 3	Block 4	Block 5
Workload Questionnaire					
Input condition closed					
Audio condition closed					
Resync both computer clocks					x
Post-Study Questionnaire	x	x	x	x	
Demographics Questionnaire	x	x	x	x	

Appendix J: Order of Trials for Each Condition

Q-NFF = QWERTY – non-feedforward

Q-FF = QWERTY – feedforward

S-NFF = softkeys – non-feedforward

S-FF = softkeys – feedforward

Trial Order for Q-NFF	Trial Order for Q-FF	Trial Order for S-NFF	Trial Order for S-FF
P>V>H9>B>E	P>V>H9>B>E	P>V>H9>B>E	P>V>H9>B>E
Y>K5>E>A	Y>K5>E>A	Y>K5>E>A	Y>K5>E>A
B>P>N>F2>L	B>P>N>F2>L	B>P>N>F2>L	B>P>N>F2>L
V>N>L>K5>E>Y	V>N>L>K5>E>Y	V>N>L>K5>E>Y	V>N>L>K5>E>Y
A>F2>N>L	B>V>N>F2>E	P>N>F2>E>Y	F2>N>L>A>B>Y
P>N>F2>E>Y	Y>E>H9>B>P	P>E>A>F2	P>N>F2>E>Y
P>L>Y>E>A>F2	V>H9>A>L>P>N	F2>N>L>A>B>Y	N>L>A>F2>E
N>F2>E>A>L	H9>Y>N>L	H9>A>B>Y	V>H9>A>L>P>N
B>Y>K5>E>A	G8>P>V>A	A>F2>L>Y>E	P>G8>V>N>B>Y
G8>P>V>N>L	N>B>E>Y>L>K5	P>G8>V>N>B	A>L>P>G8
H9>L>P>V>A	B>Y>K5>E>A	G8>P>V>A	N>F2>E>A>L
H9>A>B>Y	P>G8>V>N>B>Y	B>V>A>L>P>G8	L>P>V>A>B>G8
V>A>L>K5	A>F2>L>Y>E	N>F2>E>A>L	G8>P>V>A
G8>P>V>A	P>N>F2>E>Y	B>V>N>F2>E	P>G8>V>N
L>P>V>A>B>G8	L>P>G8>V>A>B	K5>E>Y>B	G8>P>V>N
N>B>E>Y>L>K5	K5>E>Y>B	V>A>L>K5	V>H9>L>P>N>B
E>A>L>K5	G8>P>V>N	G8>P>V>N>L	K5>E>Y>N>B>P
Y>K5>N>B	V>H9>Y>E>A	L>P>V>A>B>G8	H9>L>P>V>A
B>V>A>L>P>G8	H9>L>P>V>A	N>L>A>F2>E	N>B>E>Y>L>K5
K5>E>Y>B	L>P>V>A>B>G8	V>H9>A>L>P>N	E>A>L>K5
K5>E>Y>N>B>P	P>E>A>F2	E>H9>Y>N>B	A>F2>N>L
P>G8>V>N>B>Y	V>A>L>K5	Y>E>H9>B>P	Y>E>H9>B>P
H9>Y>N>L	N>L>A>F2>E	P>L>Y>E>A>F2	H9>Y>N>L
F2>N>L>A>B>Y	Y>K5>E>A>B>P	Y>B>P>V>H9>A	A>F2>L>Y>E
G8>P>V>N	Y>K5>N>B	H9>Y>N>L	L>P>G8>V>A>B
P>E>A>F2	P>L>Y>E>A>F2	H9>L>P>V>A	Y>B>P>V>H9>A
P>G8>V>N>B	P>G8>V>N	A>F2>N>L	E>H9>Y>N>B
V>H9>Y>E>A	V>H9>L>P>N>B	K5>E>Y>N>B>P	B>Y>K5>E>A
P>G8>V>N	A>F2>N>L	A>L>P>G8	P>E>A>F2
B>V>N>F2>E	A>L>P>G8	P>G8>V>N	Y>K5>E>A>B>P
Y>K5>E>A>B>P	P>G8>V>N>B	G8>P>V>N	V>H9>Y>E>A
Y>B>P>V>H9>A	H9>A>B>Y	E>A>L>K5	B>V>A>L>P>G8
A>L>P>G8	G8>P>V>N>L	P>G8>V>N>B>Y	P>G8>V>N>B
V>H9>A>L>P>N	K5>E>Y>N>B>P	N>B>E>Y>L>K5	K5>E>Y>B
A>F2>L>Y>E	N>F2>E>A>L	Y>K5>E>A>B>P	H9>A>B>Y
E>H9>Y>N>B	B>V>A>L>P>G8	Y>K5>N>B	V>A>L>K5
Y>E>H9>B>P	E>H9>Y>N>B	V>H9>Y>E>A	G8>P>V>N>L
V>H9>L>P>N>B	E>A>L>K5	V>H9>L>P>N>B	B>V>N>F2>E
N>L>A>F2>E	F2>N>L>A>B>Y	L>P>G8>V>A>B	Y>K5>N>B
L>P>G8>V>A>B	Y>B>P>V>H9>A	B>Y>K5>E>A	P>L>Y>E>A>F2

Appendix K: Order of Trials for Each Participant

P1, P2, P3… = Participant 1, Participant 2, Participant 3…

Block 1, Block 2, Block 3… = Condition 1, Condition 2, Condition 3…

S-NFF = softkeys – non-feedforward

S-FF = softkeys – feedforward

Q-NFF = QWERTY – non-feedforward

Q-FF = QWERTY – feedforward

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16
Block 1	Q-NFF	Q-FF	S-NFF	S-FF	W	S-FF	W	Q-NFF	Q-FF	S-NFF	S-NFF	Q-FF	Q-NFF	W	S-FF	W
Block 2	Q-FF	S-NFF	S-FF	W	Q-NFF	S-NFF	S-FF	W	Q-NFF	Q-FF	Q-FF	Q-NFF	W	S-FF	S-NFF	Q-NFF
Block 3	W	Q-NFF	Q-FF	S-NFF	S-FF	W	Q-NFF	Q-FF	S-NFF	S-FF	S-FF	S-NFF	Q-FF	Q-NFF	W	S-FF
Block 4	S-NFF	S-FF	W	Q-NFF	Q-FF	Q-FF	S-NFF	S-FF	W	Q-NFF	Q-NFF	W	S-FF	S-NFF	Q-FF	Q-F
Block 5	S-FF	W	Q-NFF	Q-FF	S-NFF	Q-NFF	Q-FF	S-NFF	S-FF	W	W	S-FF	S-NFF	Q-FF	Q-NFF	S-NFF