1-1-2014

Understanding How Age, Input Method, and Input Content Impact User Error

Angela Brooke Cannon

Understanding how age, input method, and input content impact user error

By

Angela Brooke Cannon

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Industrial Engineering
in the Department of Industrial and Systems Engineering

Mississippi State, Mississippi

December 2014

Understanding how age, input method, and input content impact user error

By

Angela Brooke Cannon

Approved:

Lesley Strawderman
(Major Professor)

Stanley F. Bullington
(Committee Member)

Reuben F. Burch, V
(Committee Member)

Daniel W. Carruth
(Committee Member)

Kari Babski-Reeves
(Graduate Coordinator)

Jason M. Keith
Interim Dean
Bagley College of Engineering

Name: Angela Brooke Cannon

Date of Degree: December 13, 2014

Institution: Mississippi State University

Major Field: Industrial Engineering

Major Professor: Lesley Strawderman

Title of Study: Understanding how age, input method, and input content impact user error

Pages in Study: 38

Candidate for Degree of Master of Science

The growing number of mobile devices used today and the increasing dependency on them in the workplace makes understanding how users interact with these devices critical. This study looks to find how different generational groups commit errors on different types of devices. Participants completed tasks consisting of word and character input on two different devices, a physical keypad and touchscreen device. The number of errors and types of error, corrected and permanent were collected for each participant. It was found that participants committed more errors when using character input and physical keypad devices but also corrected more of their errors when using them. When looking at number of errors and the amount of corrected errors, the optimal input content and input method paired combination is using word input on the key device. The results of this study can help guide industries in choosing the right devices for their users.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Text entry research has evolved over time with the growing number of mobile devices in the market. In January of 2014, 90% of American adults owned a cellphone and 58% of American adults had a smartphone (Pew Research, 2014). Text entry research has also progressed as technological advances have laid the foundation for innovative methods of analyzing text input (Kano & Read, 2009). The necessity for text entry research is driven by the need for these devices in the workforce and the technological capabilities necessary for these devices for applications in the workforce. Text entry research can help developers understand more about the speed and accuracy of a user's text input thus inspiring new text input technologies. The development of new input technologies is important not only because of the increase in mobile devices that use text input in the field but also the dependency placed on text input devices in the workforce. As technology becomes more prevalent in industry, the dependency placed on input technologies used on mobile devices can become critical. As businesses move toward increasing the use of these rugged, mobile technologies, the capabilities that the devices need become more advanced. With new technologies, the process of serving the customer becomes quicker, creating a need for the technology to be accurate and efficient.

This study will evaluate how user error changes with different input methods and input content on rugged mobile devices. Observing changes in user error allows for a better understanding of the accuracy of text input. Determining if differences exist in user error will give insight that can benefit text entry research and development by showing how error rates and types change with different devices and types of text content. This allows application developers to modify existing technologies or create new ones to increase user accuracy when inputting text.

CHAPTER II

LITERATURE REVIEW

2.1 Error Types

Accuracy and speed are the two primary evaluation metrics for text input. (MacKenzie & Soukoreff, 2002) However, determining the accuracy of text input is not as straightforward as determining input speed (Soukoreff & Mackenzie, 2001). In efforts to learn more about text input accuracy, many studies over the years have contributed to the knowledge base of typing error categorization. Classifying typing errors gives a more complete analysis of a text input evaluation compared to calculating a set error percentage to apply to all users (MacKenzie & Soukoreff, 2002). A set error percentage would not provide a clear understanding of how many errors are actually being committed specific to the study's conditions; it would merely be estimating an average percentage of errors based on historical data and applying it to all results.

Substitution, omission, and insertion are the three basic types of errors; however, as shown in Table 2.1, many researchers have expanded these categories to provide more thorough analyses of errors. It is common for methods of categorization to be developed to target a specific field of research or specific application domain (Kano & Read, 2009). Performing these error analyses can be difficult based on the number of participants and trials being completed in the experiment as well as increasing the possibility of calculation errors if being performed manually (MacKenzie & Soukoreff, 2002). The table below was developed by Kano & Read to summarize typing error categorization methods from 1945 to 2007 as they developed methods to solve character level insertion ambiguities. Kano & Read noted that as the number of categories for classifying error type grew, as seen in Table 2.1, the possibility of an error falling into one of the multiple categories also increased. Their research focused on eliminating the ambiguities associated with the possibility of errors being classified in more than one category.

Table 2.1 Summary of Typing Error Categorization Methods from 1945 to 2007

	White (1936)	Dvorak, A., Merrick, N. L., Dealey, W. L., and Ford, G. C. (1936)	MacNeilage (1964)	Gentner (1983)	Logan (1999)	Wobbrock & Myers (2006)	Kano, A., Read, J. C., Mackenzie, I. S., & Dix, A. (2007)
Insertion	Inserted strokes, Double strokes, Syllable division, Repeating words		Interpolation error	Insertion, Misstrokes	Immediate, preservation, Space bar, Separation, Character separation, Letter sequence, Error habits, Home-letter intrusion	Insertion	CT-Mu, DL, DS, ExE, IF, IL, IS, ISy, NT-MA, IW-A, IW-U, DW, DP, EE
Omission	Letter omission, Omitting words	Omission, Omitting words	Omission errors	Omission	Letter, Syllable, Word, Space bar	Omission	OL, OS, OW, OP
Substitution	Substitution, Transposition, Capitalization, Word substitution, Adjacent letter substitution	Substitution, Adjacent errors, Copy reading, Transposition, Transposed doubling	Horizontal, Vertical, Diagonal, Reversal, Anticipation, Phonemic, Type errors, Contralateral	Substitution, Transposition, Interchange, Migration, Doubling, Alternating	Remote, Horizontal, Vertical, Number sub, Homologous, Hand position, Transposition (1 finger, 2 finger, 2 hands), Interchange, Migration, Alternation, Doubling, Another word, Perseveration	Substitution	AE, CE, CT-S, DE, IE, NT-S, SL, TE, ME, SW-A, SW-U, SP
Other	Spacing		Equivocal, Multiple classification, Unclassifiable		Spelling		U, ExE, CNE(error)

(Adapted from Kano and Read 2009)

2.2 Input Method

Two of the most common input methods for handheld devices are keypads and touchscreens. While both devices are used for text input, keypads have physical key affordances which have been proved to be vital when typing (Paek & Hsu, 2011). When using a touchscreen, users are more prone to error without the affordances a keypad provides (Rudchenko, Paek, & Badger, n.d.), and as a result perform text entry worse than on a keyboard. A 1990 study asked participants to enter dates and airline flights on a keyboard and a touchscreen (both without autocomplete features), and when looking at input speed found that users of touchscreens took 73% longer when inputting text content in comparison to a keyboard; however the ubiquity of devices in our daily lives could change these values today. (Gould, Greene, Boies, Meluson, & Rasamny, 1990). In a study where participants performed tasks on an ATM with a numeric physical keypad and one with a touchscreen keypad, the input speed of older adults on a touchscreen was found to be faster than a physical keypad (Chung, Kim, Na, & Lee, 2010; Rogers, Fisk, McLaughlin, & Pak, 2005).

There is a lack of significant studies on keyboard-based text input alone. The number of studies began to decrease with the increase in popularity of graphical user interfaces. In more recent years, with the growth in popularity of mobile devices, text input studies have become more prominent; however, finding studies concentrated on the keyboard are still rare (Kano et al., 2007).

2.3 Input Method and Age

There have been a number of studies that seek to find age-related differences when using direct and indirect input devices (Chung et al., 2010; Mclaughlin, Rogers, & Fisk, 2012; Pak, McLaughlin, Lin, Rogers, & Fisk, 2002; Rogers et al., 2005). A device is defined as indirect or direct by its manner of input and output of data (Mclaughlin et al., 2012). A device where input is completed on the output screen, such as a touchscreen, is considered a direct device. Devices that separate the method of input from the output screen, such as a keyboard or mouse, are considered indirect devices.

Pak, et al. (2002) focused their study on determining which device resulted in better performance for younger adults and older adults, average ages 19.95 and 58.45 respectively. The results yielded the completion of tasks more quickly for both age groups on the touchscreen device. The study reveals the implications of how the type of task can impact the input speed of specific age groups performing on different devices. Rogers, et al. (2005) also concluded that age and task type can impact which type of device is optimal for a task. When using indirect devices, more attention is demanded of the users due to required translation between the input and output (Rogers et al., 2005). Relevant to age, older adults tend to exhibit difficulty when divided attention is required (Mclaughlin et al., 2012).

Chung et al. (2010) investigated the impact of age on direct and indirect numeric keypad types, yielding results that showed older adults again performing faster when using a touchscreen but also making more errors compared to the physical keypad. Chung attributes this increase in error to the lack of tactile feedback on a touchscreen. It seems most literature concludes that the touchscreen allows both older and younger groups to perform faster and would suggest it as the optimal device but one cannot assume this is always true. Rogers et al. (2005) found that, while for younger adults the type of task can often determine the optimal input device, older adults do not return a steady pattern. Rogers concludes that with older adults, individual differences may be the indicator impacting the optimal device for a specific task.

CHAPTER III

INITIAL STUDY

In a previous study completed in the Mississippi State University Department of Industrial and Systems Engineering, two different generational groups used two different types of ruggedized handheld devices to determine which input type is best for each group (Burch, 2014). Of the two generational groups in Burch's study, one group was composed of Baby Boomer- aged employees (between 50 and 68) and the other included Gamer-aged employees (between 18 and 35). Learning more about the differences of these two generations has relevance as organizations are being impacted by the retirement of the Baby Boomer generation and hiring more of the incoming Gamer generation (Burch and Strawderman, 2014). Of the Boomer participants in the study, 85% owned cell or smart phones while 100% of gamer participants owned one. Each group manually entered the contents of a task list, completing a total of 30 tasks. To complete a task, participants entered an airplane container name, composed of alpha and numeric characters, and a comment, composed of words. Completion time for the tasks was measured and the number of errors in each entry was counted. The experiment found that there was a significant difference in the manual entry time between the two generational groups and between the two different rugged device types. The error differences weren't significant but they were worthy of further study.

CHAPTER IV

OBJECTIVES AND HYPOTHESES

4.1 Research Objective

This study serves as a continuation of the previous experiment to learn more about user error (Burch, 2014). The main objective of this study is to determine how the user's error rate is affected by age, input content, and input method. This will be accomplished through observation of the recorded task performance on the ruggedized handheld devices.

4.2 Hypotheses

The literature review and objective in the previous sections contribute to form the following hypotheses to be tested in this experiment.

Hypothesis 1: There will be more total errors made in word inputs than character inputs.

There is a lack of research in comparing types of text input content. The growing number of mobile device users and the increasing number of text messages sent worldwide proves the use of handheld text input devices to be a familiar territory for most users. Today, 90% of American adults own cell phones and 81% of cell phone owners use their cell phone to send or receive text messages (Pew Research, 2014). Many of these devices also feature correction software for users when typing word inputs. With participant expectancy of the corrective features to fix typing mistakes, it is predicted that when typing words participants will commit more errors. On most keyboards typing combinations of alpha and numeric characters requires more concentration since it is an unfamiliar sequence of characters, unlike words which are recognizable sequences used daily in text messages, email, and other word processing applications. This also supports the hypothesis as it is predicted since participants will likely concentrate more on the input of character strings because of their unfamiliarity with the context, which will result in fewer errors.

Hypothesis 2: When comparing errors made on the paired combinations of input content and input method, the greatest number of errors will be made when entering word input on the touchscreen.

Supported by the literature (Chung et al., 2010; Rudchenko et al., n.d.), it is predicted that performance on the touchscreen will result in more errors than on the physical keypad. The research attributes this increase in errors on touchscreens to its lack of affordances. In agreement with hypothesis 1 above, this would result in the most errors when inputting word content on a touchscreen.

Hypothesis 3: When performing word input on touchscreen, there will be more permanent errors than corrected errors.

As stated in the previous hypothesis, there will be more errors on the touchscreen when performing word input; however, among those errors it is predicted that there will be more permanent errors than corrected errors. It is expected that without corrective software, users will be less attentive to errors being made as they are accustomed to their devices fixing their errors for them. This will results in an increase of permanent errors.

Hypothesis 4: When comparing age and input content, Gamers performing word input will have the highest number of total errors.

With the Gamer generation growing up immersed in technology, it is likely that they are less attentive to their actions when using the devices. Word input is a daily task for Gamers, resulting in Gamers going through the motions of typing words without noticing their errors.

CHAPTER V

METHODS

5.1 Initial Experiment Overview

As mentioned above, this study is a continuation of the study conducted by Burch (2014). The following are the methods used in the initial study.

5.1.1 Participants

Data was collected from forty participants, including twenty people from the Gamer generation and twenty people from the Boomer generation. Each participant was a regular user of a ruggedized keypad device and experienced with a touchscreen device. Of the participants, 85% were male. The higher proportion of male participants was due to the lack of availability of female participants at the selected locations.

5.1.2 Procedures

The experimental procedure was composed of three elements: 1) a questionnaire about demographics and device preference of the participant, 2) recorded data entry on the first rugged handheld device, 3) recorded entry on the second rugged handheld device.

At the beginning of the procedure, prior to any data entry, each participant was asked a brief series of questions. These questions established whether the participant was a member of the Gamer generation or the Boomer generation, their age, and their work job function. Participants were also asked if they currently owned a cell phone or smartphone and about their experience level with consumer-grade handheld devices that have a keypad or touchscreen. Participants were then asked to confirm their years of ruggedized device experience.

The type of device used for the first data entry segment was selected at random. For the purpose of counter balancing, ten participants from each age group started with the keypad and ten started with the touchscreen. The two ruggedized handheld devices used were the Motorola MC9500 and the Motorola ETI. These two devices had like characteristics, such as key surface area, height, and weight. Participants were experienced with theMC9500 devices and used the key device on a daily basis.

Each participant was asked to manually enter the contents of the task list into a word processor application installed on the ruggedized device. For each device, participants were provided a task list with 30 airplane container names and 30 associated comments. The 30 container names were composed of 10 characters, alpha and numeric, and the associated comments were composed of words, 19-24 characters. Below in Figure 5.1 is an excerpt from the task list.

DATA ENTRY TASKS (B-5)

Device 1 (List 20)

#	Container	Comments
1	AKE53190FX	contains postal packages
2	AVE47013FX	saturday document bag
3	SAA76459FX	careful dangerous goods
4	SAX23174FX	fyi hazardous materials
5	AYX67812FX	dry ice multi piece ship

Figure 5.1 Screenshot of Excerpt of Data Entry Task List

After each line, one carriage return was required. After completing each entry set of container name and comment, an additional carriage return was required. Notes were provided on the task worksheet reminding the participant the rules regarding carriage returns and data entry format.

Participants were told to enter the data as quickly as possible but still try to be as accurate as possible. Fixing mistakes was encouraged. All participants' actions to complete the tasks, including hand movement and data entry on the screen, were video captured using a Mr. Tappy filming kit and a Swann DIY Security Camera with built in microphone attached to the device. Each participant completed a total of 30 tasks on each device but only the middle 20 tasks were used for data analysis. With possible learning curves for the typing tasks, removing the first and last five tasks were expected to eliminate skewed results associated with the user's adaptation to the device as well as fatigue.

5.2 Continuation Study

5.2.1 Independent Variables

5.2.1.1 Age

Participants were categorized into two age groups, the Gamers and the Boomers, based on their age at the time of the study. Gamers consist of those in the work force that were born between 1995 and 1979 and boomers consist of those in the workforce born between 1945 and 1963.

5.2.1.2 Input Method

There were two input methods in this experiment: physical key input and touchscreen input. Two devices (Figure 5.2) were chosen based on the current use of the ruggedized handheld devices with a keypad and the future of devices being primarily touchscreen. Input method was a within-subjects variable since all users performed tasks with both devices.

Figure 5.2 The Motorola MC9500 [left] and Motorola ET1 [right]

5.2.1.3 Input Content

Participants entered two different types of input content when using the rugged handheld devices. The first input content type was character. Character input included strands of characters, alpha and numeric. The second input content type was word. Participants entered strings of complete words; however, this did not necessarily include full sentences. These two input content types simulated the input content employees enter on the devices in the workplace. Input content is a within-subjects variable since users input both word and character content during the experimental task. Below is an example of each input content type.

Character input content: AKE01241FX

Word input content: hold shipment at station

Due to the unequal number of characters used in the data entry tasks for the two types of input content, the number of errors committed was divided by the maximum possible number of word or character errors for that data entry. This calculated an error percentage for each data entry to create an equal probability of erring when entering either of the two input content types. To calculate the maximum number of errors, the number of possible errors for the data entry was calculated. For character input, each entry always consisted of 10 characters. For word input, the longest possible entry was 24 characters in length. The maximum strand length was used to ensure all possible errors were accounted for. These percentage values were used to allow equal comparisons of all variables. Below is an example calculation for the method described.

Example word input strand: hold shipment at station

|h|o|l|d| |s|h|i|p|m|e|n|t| |a|t| |s|t|a|t|i|o|n|
24 characters in length

Maximum Number of Possible Errors:
(24 characters maximum per task)(20 tasks per participant for tasks 6-25) = 480 possible errors

Word Errors for Participant A3 on Key Device Tasks 6-25:
7 word errors/(480 maximum possible errors for tasks 6-25) = 0.0146 =1.46 %

5.2.2 Dependent Variables

5.2.2.1 Number of Errors

The number of errors, corrected and permanent, were counted for each user. Errors were sorted into the following categories: insertion, omission, sequence, and substitution. All of these categories serve as a method to define how to count each error individually. Insertion, omission, and substitution were all categories noted in the literature review. While sequence was not referenced, it has been added to this study to better distinguish the number of errors committed. For example, if the user typed the word “Dnevet” instead of “Denver”, the number of errors would be two. The user committed a substitution error where the “r” should be placed at the end of the sentence and a sequence error by switching the “e” and the “n.” If sequence was not a category, the switch of two letters may be considered two substitution errors when it is most likely that the user only committed one error by switching the two letters, not substituting each. The classification of errors was not analyzed, but will be reported in the descriptive statistics of the study.

5.2.2.2 Error Types

In this experiment there were two types of error. The first type of error was a corrected error. A corrected error was when a user input the wrong character but went back to their error to input the correct character. The second type of error was a permanent error. A permanent error was an error that the user made but did not go back to correct. This was measured through observing the participants’ performance, determining whether the participants committed errors, and if they attempted to resolve those errors or not. If at any point throughout the experiment the user attempted to resolve the error and corrected it, it was considered a corrected error. If the error remained uncorrected until the end of the experiment, it was considered permanent. When participants erred, attempted to resolve it, but again entered an incorrect character, they were considered to have committed two errors in total.

5.2.3 Data Extraction

Using the video captured from the previous experiment, participants' task performance was reviewed to count the errors and classify the error types within the word and character inputs. An example screenshot from the video is shown in Figure 5.3. The classification of the errors into the four groups was only used for counting the number of errors. Errors were sorted into three groups for tasks 1-5, 6-25, and 26-30.

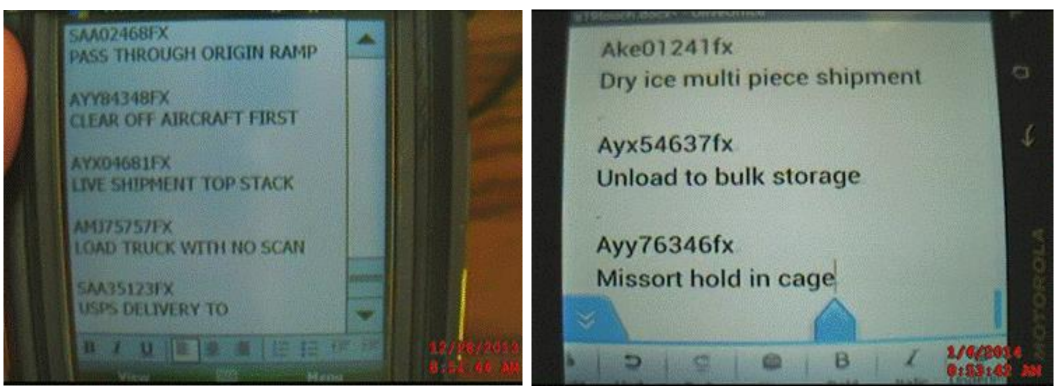

Figure 5.3 Screenshot of video captured for both input methods, key (left) and touch (right)

5.2.4 Analysis

IBM SPSS Statistics version 21 (SPSS IBM, New York, U.S.A) was used to analyze the data collected for this study. An alpha value of 0.05, or confidence level of

95%, was used to evaluate levels of significance. Appropriate descriptive statistics were calculated. Inferential analysis included two-sample t-test (hypothesis 1), factorial ANOVA (hypothesis 2 and 4), and one-sample t-test (hypothesis 3). Data were tested for appropriate analysis assumptions (e.g. normality for ANOVA) prior to analysis. Appropriate data transformations were made as needed or a suitable non-parametric analysis would have been used if necessary.

CHAPTER VI

RESULTS

6.1 Descriptive Statistics

6.1.1 Number of Errors by Error Category

Table 6.1 includes the number of insertion, omission, substitution, and sequence errors committed for tasks 6-25. The error categories in the table below were used to determine the number of errors for each variable.

Table 6.1 Total Number of Errors by Error Type Categorization Summary Tasks 6-25

	INSERTION	OMISSION	SUBSTITUTION	SEQUENCE
OVERALL				
	234	487	1418	61
GENERATION				
BOOMER	113	268	693	24
GAMER	121	219	725	37
INPUT CONTENT				
WORD	154	356	897	40
CHARACTER	80	131	521	21
INPUT METHOD				
KEY	123	191	652	28
TOUCH	111	296	766	33
ERROR TYPE				
PERMANENT	103	329	472	38
CORRECTED	131	158	946	23

6.1.2 Number of Errors

Table 6.2 provides the mean, standard deviation, and range of the number of errors for generation, input content, input method, and error type for the data entry tasks. Table 6.2 includes the data for tasks 6-25.

Table 6.2 Number of Errors Summary Tasks 6-25

	MEAN	STANDARD DEVIATION	MIN	MAX
GENERATION				
BOOMER	27.5	28.6	7	235
GAMER	27.6	18.3	12	100
INPUT CONTENT				
WORD	36.2	27.1	6	157
CHARACTER	18.8	14.6	1	78
INPUT METHOD				
KEY	24.9	27.1	2	147
TOUCH	30.2	20.1	5	88
ERROR TYPE				
PERMANENT	23.6	28.2	1	144
CORRECTED	31.5	19.0	5	91

6.1.3 Percentage of Errors by Error Category

Table 6.3 includes the percentage of insertion, omission, substitution, and sequence errors committed for tasks 6-25. The percentage values were calculated by taking an average of the percentage values that were calculated the same method as described previous. While the values in Table 6.1 and Table 6.3 were not analyzed in this study, they served a purpose calculate the number of errors for each variable.

Table 6.3 Average Percentage of Errors by Error Type Categorization Summary Tasks 6-25

	INSERTION	OMISSION	SUBSTITUTION	SEQUENCE
OVERALL				
	1.8%	3.5%	11.2%	0.5%
GENERATION				
BOOMER	1.7%	3.7%	10.6%	0.3%
GAMER	1.9%	3.2%	11.8%	0.6%
INPUT CONTENT				
WORD	0.8%	1.9%	4.7%	0.2%
CHARACTER	1.0%	1.6%	6.5%	0.3%
INPUT METHOD				
KEY	1.0%	1.4%	5.9%	0.2%
TOUCH	0.8%	2.1%	5.3%	0.2%
ERROR TYPE				
PERMANENT	0.8%	2.3%	3.4%	0.3%
CORRECTED	1.0%	1.2%	7.8%	0.2%

6.1.4 Percentage of Errors

Table 6.4 provides percentage values of the mean, standard deviation, and range of the percentage of errors for generation, input content, input method, and error type for the data entry tasks. These percentage values were calculated using the methods described previously.

Table 6.4 Percentage of Errors Summary Tasks 6-25

	MEAN	STANDARD DEVIATION	MIN	MAX
GENERATION				
BOOMER	16.3%	14.8%	1.8%	71.7%
GAMER	17.6%	8.2%	3.1%	31.8%
INPUT CONTENT				
WORD	7.5%	5.6%	1.3%	32.7%
CHARACTER	9.4%	7.3%	0.5%	39.0%
INPUT METHOD				
KEY	8.5%	9.1%	0.7%	49.0%
TOUCH	8.5%	5.8%	1.0%	24.6%
ERROR TYPE				
PERMANENT	3.5%	4.2%	0.1%	21.2%
CORRECTED	4.7%	2.8%	0.7%	13.4%

Figure 6.1 provides a summary of the percentage of all errors when categorized based on generation, input content, input method, and error type. The data in Figure 6.1 is only for data entry tasks 6-25.

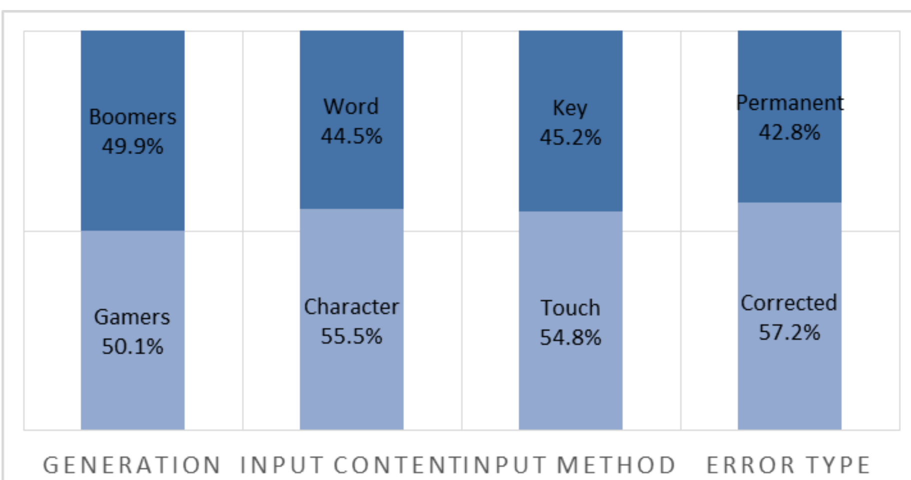

Figure 6.1 Total Percentage of Errors

Regarding total number of errors, both generations totaled close to the same number of errors; the Gamer generation made only four more errors (0.18%) than the Boomer generation overall, as seen in Figure 6.1. Input content, input method, and error type all had a greater difference compared to generation.

Regarding individual participants' performance with each variable, when entering input content, twenty-three participants made more errors when entering character input compared to word. Of these 23 participants, 15 were from the Gamer generation and eight were Boomers. The majority of participants from each generational group, 14 Gamer participants and 13 Boomer participants, made more errors on the touch device than the key device. Regarding error type, 17 Gamer participants had more corrected errors than permanent while this was only true for 12 Boomer participants.

6.1.5 Percentage of Errors by Variable

The tables below include the percentage of average errors broken down by each of the variables. Table 6.5, Table 6.6, Table 6.7, and Table 6.8 provide a more detailed look at how each of the variables was impacted by one another. These values were calculated by taking an average of the divided number of errors by the possible number of errors for each participant's data entry task.

Table 6.5 Percentage Average Error by Generation

	GAMERS	BOOMERS
GENERATION		
OVERALL	17.6%	16.3%
INPUT CONTENT		
WORD	7.1%	8.0%
CHARACTER	10.5%	8.4%
INPUT METHOD		
KEY	7.8%	9.2%
TOUCH	9.8%	7.0%
ERROR TYPE		
CORRECTED	12.5%	7.9%
PERMANENT	5.0%	8.4%

Table 6.6 Percentage Average Error by Input Method

	KEY	TOUCH
INPUT METHOD		
OVERALL	8.5%	8.5%
GENERATION		
GAMERS	7.8%	9.8%
BOOMERS	9.2%	7.0%
INPUT CONTENT		
WORD	2.8%	4.7%
CHARACTER	5.7%	3.7%
ERROR TYPE		
CORRECTED	5.3%	4.9%
PERMANENT	3.2%	3.5%

Table 6.7 Percentage Average Error by Input Content

	WORD	CHARACTER
INPUT CONTENT		
OVERALL	7.5%	9.4%
GENERATION		
GAMERS	7.1%	10.5%
BOOMERS	8.0%	8.4%
INPUT METHOD		
KEY	2.8%	5.6%
TOUCH	4.7%	3.7%
ERROR TYPE		
CORRECTED	3.9%	6.3%
PERMANENT	3.6%	3.1%

Table 6.8 Percentage Average Error by Error Type

	CORRECTED	PERMANENT
ERROR TYPE		
OVERALL	10.3%	6.7%
GENERATION		
GAMERS	12.5%	5.0%
BOOMERS	7.9%	8.4%
INPUT METHOD		
KEY	5.2%	3.2%
TOUCH	4.9%	3.5%
INPUT CONTENT		
WORD	3.9%	3.6%
CHARACTER	6.3%	3.1%

6.2 Inferential Statistics

Hypothesis one stated there would be more total errors made in word inputs than character inputs. After calculating new values to create an equal probability of committing an error among the two input content types, statistical testing revealed that the difference between the means of word input and character input were not significant, t(158)=1.34, p=0.182. Below, Figure 6.2 shows the average error percentage for character and word input.

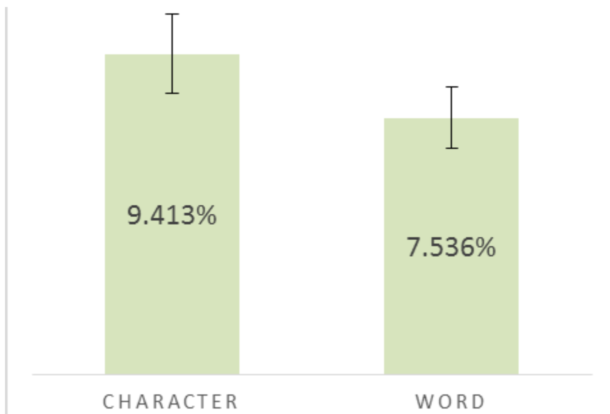

Table 6.9 Error Percentage of Input Content

Hypothesis two stated that performing word input on a touch device would result in the greatest number of errors. While this hypothesis was not proven correct by the observed data, the statistical analysis did find this model to be significant, F(3,159)=3.289, p=0.022. The interaction between input method and input content was also found to be significant, F(1,159)=8.001, p=0.005. Figure 6.3 below shows the average percentage for each of these combinations.

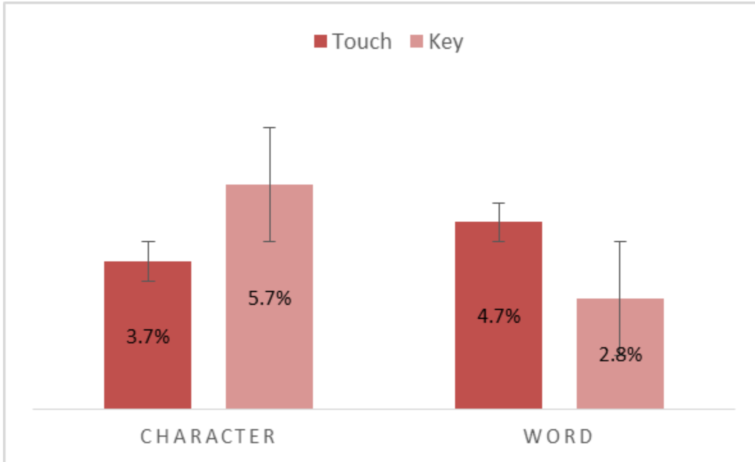

Figure 6.2 Percentage of Error by Input Content and Input Method

Hypothesis three stated that when performing word input on a touch device, there would be more permanent errors than corrected errors. Statistical testing showed that the difference between the number of corrected and permanent errors among the word input content and touch input method device was not significant, t(78)= -0.97, p=0.334. Figure 6.4 below presents the observed data for hypothesis three.

While hypothesis three only stated the combination of word input on a touch device, after taking note of the observed data, further statistical testing was done for the remaining input content and input method combinations. Of the three remaining combinations, the character input on a key device was statistically significant, t(78)=2.314. p= 0.023. Character input on a touch device, t(78)=0.210, p= 0.040, and word input on a key device, t(78)=0.058, p=0.833, were both insignificant.

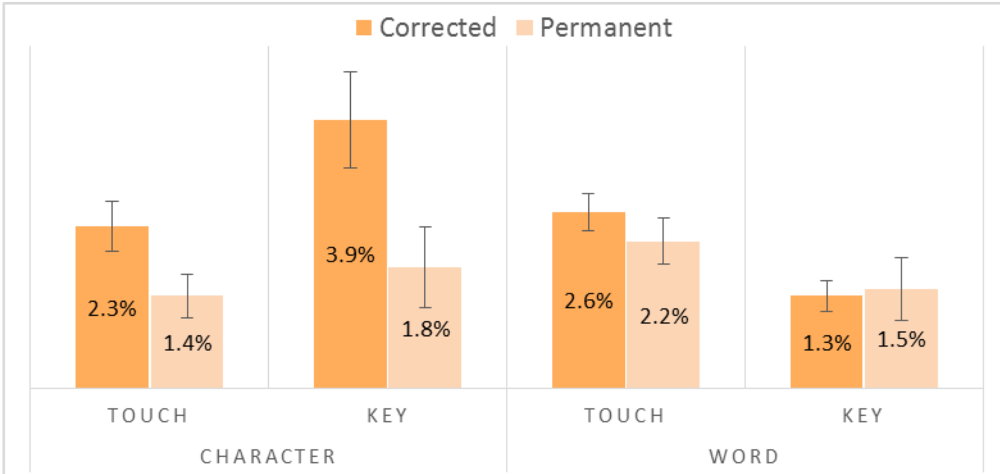

Figure 6.3 Percentage of Error by Error Type, Input Method, and Input Content

Hypothesis four stated that when comparing age and input content, Gamers performing word input would have the highest number of total errors. Statistical tests found the model to be insignificant, F(3,159)=1.040, p=0.377. Figure 6.4 depicts the observed data for the fourth hypothesis.

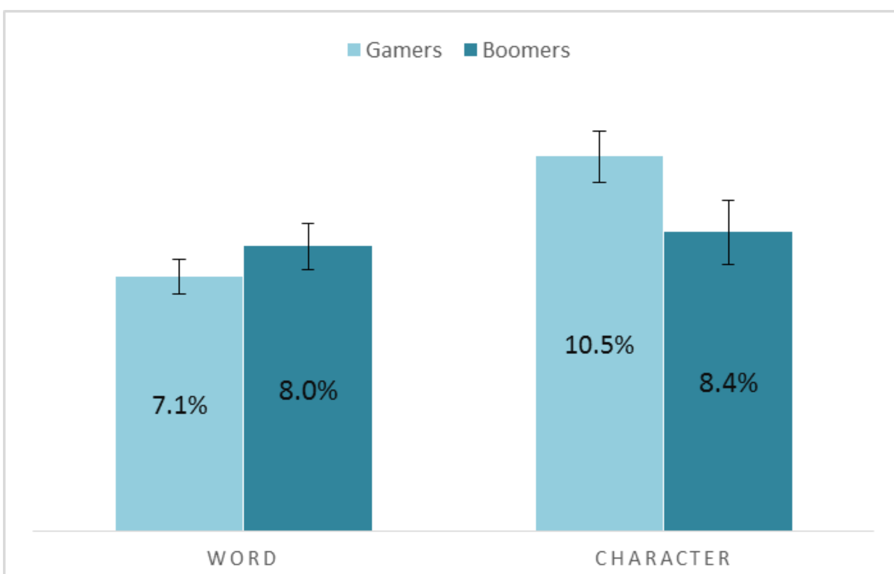

Figure 6.4 Percentage of Errors by Input Content and Generation

CHAPTER VII

DISCUSSION

The goal of this study was to learn more about the relationship between generational groups and the way they use handheld ruggedized devices. The four hypotheses aimed to target more specific areas to show how different elements impact device usage.

When learning to type, memorizing the location of the keys is the first step, but as one becomes accustomed to the keyboard, the patterns of words become familiar. Character input content in this study was a ten-character strand of alpha and numeric values that does not contain any words. This could have led to an increase in attention to the output of the task due to the inability of the participant to use familiar typing patterns, leading the participant to become aware of errors made. However, if the increase in attention could be a reason why there were more corrected errors, why would this not have lessened the total number of errors when entering character input on the key device? While the methods of data collection in this study prevent this possibility from being evident, the number of character errors could be attributed to reading errors of the participants. Similar to remembering passwords, users tend to struggle with remembering random characters (Topkara, 2007). Reading and remembering words that form a fragment of a sentence to translate to a word processor is simpler than a random ten-character alpha numeric strand. This lack of tangible words or sentences could have led to an increase in errors when translating the character strands from the task list to the word processor.

Results concerning the second hypothesis showed that the greatest number of errors was made when entering character input on the key device and second greatest, word input on the touch device. The absence of correction software in the study led to an underestimation of the number of errors on touch devices. As most users of mobile devices and tablets are accustomed to correction software, it was expected that errors would often be glanced over in anticipation of device auto-correction. The hypothesis expected the increase in errors due to lack of auto-correction software to be amplified when using word input on touch device, as word input on touch devices are so ubiquitous in daily activities; however, this was not the case. Devices with a lack of auto-corrective software may not impact user performance. While auto-corrective software is beneficial, its necessity isn't evident from the results of hypothesis two.

When testing hypothesis two, the model, as well as the interaction between input method and input content, was found to be significant. This is an important finding from the results of this study. This interaction highlights the possible effects of combining certain input contents and input methods. Each hypothesis resulted in the greatest number of errors when using character input content, key input method, and, for two hypotheses, a combination of both. This could imply that there is a higher probability of committing errors when entering character input on the key device. A previous study (Chung 2010) saw an increase in errors when using the touchscreen; however, this didn't seem to be the case for this study, as the higher number of errors seemed to be associated with the key input device. Chung's reasoning that an increase in errors was due to the lack of affordances on touchscreen devices did not duplicate in findings for this study. Furthermore, incorporating the results of hypothesis three to this finding brings an intriguing result to light. Hypothesis three found that the highest number of corrected errors was when entering character input on the key device. So not only was the highest number of errors committed when using the character/key combination of input method and input content, but it also resulted in the most corrected errors. But when comparing the four input content and input method combinations, determining the optimal combination cannot be solely based on the average percentage of corrected errors.

Since a high number of errors is not an ideal scenario in the workplace, determining the optimal paired combination of input method and input content must incorporate the total percentage of errors and corrected percentage of errors. When looking at the average total percentage of errors in Figure 6.3, the word/key combination is the smallest. This low number of errors overall is ideal in the workforce. The lowest average of total errors insinuates fewer opportunities for users to make any type of error, corrected or permanent. With these characteristics, the word input content using the key input method is an optimal combination.

Hypothesis four sought to find a difference among the number of errors when looking at input content like hypothesis one but added an additional variable, generation. It was predicted that the gamer generation would be most impacted by the lack of correction software when using word input due to their generation being highly accustomed to word input through technologies like text messaging; however, this was not verified in the analysis. Determining the optimal device for each age group cannot be shown from the analysis performed for hypothesis four. While determining the optimal device cannot be determined from this study's error analysis, incorporating input speed could give an alternate perspective. In the initial study conducted by Burch, input speeds were collected for each participant (Burch 2014). Among the two input methods, all users took significantly less time when completing the tasks on the touch device. Among the two age groups, Gamers take significantly less time than Boomers with both input methods, key and touch. While the error analysis does not show a significant difference among the two generational groups when committing errors, the input speeds can help choose an optimal device based on the user. When implementing new devices into the workforce, choosing touch devices will allow both user groups to perform tasks quicker.

. In the literature, the advantage of error categorization was to eliminate the ambiguity in applying a base error rate to all data entry tasks. When looking at the error rates for participants, applying a base error rate in similar tasks would be a poor assumption. The range of error rates for participants in this study portrays the differences in how participants commit errors. The high standard deviation for average percentage of errors among the Gamers (9%) and the Boomers (14%) displays the range of error rates committed in the study and highlights the necessity of calculating errors without the use of a baseline error rate.

While collecting data it was observed that, when entering character input, there seemed to be a "chunking" issue that could have impacted users' perception of the tasks and their performance of the data entry tasks. As the task list was created to model familiar workplace tasks for the, it was not an issue for learning more about how employees from this organization would be impacted but could be a short coming in assuming that others outside of the organization would behave the same. As stated previously, character input content was a ten character strand of alpha and numeric characters; however, within that ten character strand, there were three significant "chunks" of characters. The first three characters were always a combination of three letters, the next five characters were always a random combination of five numeric characters, and the last two characters were always "FX." How this chunking might have impacted the study is unknown but it can be viewed as a shortcoming of the experiment.

CHAPTER VIII

CONCLUSION

The results of this study proved different some expectations of how handheld ruggedized handheld devices are used. The significant interaction between input method and input content and the result of a greater number of errors when using character input content and key input method could suggest poorer performance when used on ruggedized handheld devices. The low average total number of errors using the word input content and key input method combinations gives an optimal combination for accurate output by allowing fewer chances for error.

While there are no other significant results or interactions, there are multiple notable results. Regarding generational groups, there was a notable underestimation of the boomer generation's ability to use more current technologies like the touchscreen device and an overestimation of the gamer generation's inattentiveness to the task. Both groups made the same amount of errors but when adding in the factor of input speed, both user groups performing significantly faster on touch devices.. For organizations working to implement ruggedized handheld devices in the workplace, dependent on the use of the ruggedized handheld device, an organization could determine which device would be the best fit based on future users and what level of quality of the output is expected from the device. Combining the input speed results from the initial study and the significance between input method and input content, using touch devices would result in better performance overall. Both generational groups perform faster on touch devices and of the input method and input content combinations, touch devices resulted in low errors when entering character input. While word input was not the lowest on the touch device, it was not the least favorable combination of the highest number of errors.

CHAPTER IX

FUTURE WORK

To learn more about user error, future research should seek to understand more about error type and the user's performance. If users are permitted to correct their errors, research should look to understand more about why they are leaving them uncorrected. Surveying users on their motives during or after the experiment may help discover reasons why these errors are being left uncorrected. Learning more about error type can help develop new devices that reduce user error rather than depending on corrective software as most word processing systems do today.

In the results above, the number of errors in each error category collected in the data extraction was reported. While these values were not statistically analyzed for this experiment, further research to analyzing these results could also reveal more about how errors are committed. Additionally in the literature review, Table 2.1 displays an even more in-depth way to analyze error types that the basic three categories used in this study. Analyzing the errors committed in this study on a deeper level may teach us more about how errors are committed. This could also reveal more about why users are committing these errors from the more elaborate categorization. Placing a value on these different types of error could provide management with more background on how employees should use these devices and which devices best suit their workforce.

REFERENCES

Burch, R. (2014). *Understanding the advantages gamers bring to the workforce and what their skillset means for the future of handheld scanning technology in large industrial organizations.* Dissertation, Mississippi State, MS: Mississippi State University.

Burch, R. & Strawderman, L. (in press). Leveraging Generational Differences to Reduce Knowledge Transfer and Retention Issues in Public Administration. *Public Administration Research*.

Chung, M. K., Kim, D., Na, S., & Lee, D. (2010). Usability evaluation of numeric entry tasks on keypad type and age. *International Journal of Industrial Ergonomics*, *40*(1), 97–105.

Dvorak, A., Merrick, N. L., Dealey, W. L., and Ford, G. C. (1936), Typewriting Behaviour. 1936: American Book Company, USA

Gentner, D. R. (1983), The Acqusition of Typewriting Skill. Acta Psychologica, 1983. 54: p. 233-248.Gould, J. D., Greene, S. L., Boies, S. J., Meluson, A., & Rasamny, M. (1990). Using a touchscreen for simple tasks. *Interacting with Computers*, *2*(1), 59–74.

Hoggan, E., Brewster, S. a., & Johnston, J. (2008). Investigating the effectiveness of tactile feedback for mobile touchscreens. *Proceeding of the Twenty-Sixth Annual CHI Conference on Human Factors in Computing Systems - CHI '08*, 1573.

IBM Corp. Released 2012. IBM SPSS Statistics for Windows, Version 21.0. Armonk, NY: IBM Corp.

Kano, A., & Read, J. C. (2009). Text Input Error Categorisation : Solving Character Level Insertion Ambiguities Using Zero Time Analysis, 293–302.

Kano, A., Read, J. C., Mackenzie, I. S., & Dix, A. (2007). ExpECT : An Expanded Error Categorisation Method for Text Input.

Logan, F. A. (1999), Errors in Copy Typewriting. Journal of Experimental Psychology: Human Perception and Performance, 1999. 25(6): p. 1760- 1773.

MacKenzie, I. S., & Soukoreff, R. W. (2002). Text Entry for Mobile Computing: Models and Methods, Theory and Practice. *Human-Computer Interaction*, *17*(2), 147–198.

MacNeilage, P. F. (1964), Typing Errors as Clues to Serial Ordering Mechanisms in Language Behaviour. Language and Speech, 1964. 7: p. 1144-1159.

Mclaughlin, A. C., Rogers, W. A., & Fisk, A. D. (2012). Using Direct and Indirect Input Devices: Attention Demands and Age-Related Differences. *ACM Transactions on Computer-Human Interaction*, *16*(1), 1–15.

"Mobile Technology Fact Sheet." *Pew Research Internet Project*. Pew Research, n.d. Web. 23 Sept. 2014.

Paek, T., & Hsu, B. P. (2011). Sampling Representative Phrase Sets for Text Entry Experiments : A Procedure and Public Resource. Microsoft Research.

Pak, R., McLaughlin, a. C., Lin, C.-C., Rogers, W. a., & Fisk, a. D. (2002). An Age-Related Comparison of a Touchscreen and a Novel Input Device. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *46*(2), 189–192.

Rogers, W. a., Fisk, A. D., McLaughlin, A. C., & Pak, R. (2005). Touch a Screen or Turn a Knob: Choosing the Best Device for the Job. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *47*(2), 271–288.

Rudchenko, D., Paek, T., & Badger, E. (n.d.). *Text Text Revolution : A Game that Improves Text Entry on Mobile Touchscreen Keyboards*.

Soukoreff, R. W., & Mackenzie, I. S. (2001). Measuring Errors in Text Entry Tasks : An Application of the Levenshtein String Distance Statistic, 319–320.

Topkara, Umut, Mikhail J. Atallah, and Mercan Topkara. "Passwords Decay, Words Endure: Secure and Re-usable Multiple Password Mnemonics."SAC '07 Proceedings of the 2007 ACM Symposium on Applied Computing (2007): 292-99. Web.

White, W. T. (1945), Typing for Accuracy. 1945: H. M. Rowe Company, Balitimore & Chicago, USA.

Wobbrock, J. O., & Myers, B. a. (2006). Analyzing the input stream for character- level errors in unconstrained text entry evaluations. *ACM Transactions on Computer-Human Interaction*, *13*(4), 458–489.