Data Entry Accuracy and Comfort with the Number Pad

Shatoyia S. Burns, Garrett J. Farnes, Heidi A. Johanson, Brian Brehman, and Kimberly A. Barchard
University of Nevada, Las Vegas

UNLV
UNIVERSITY OF NEVADA LAS VEGAS

ABSTRACT

This study investigated the relationship between comfort with the number pad and accuracy on a data entry task. Participants indicated how comfortable they were using the number pad without looking at the keys. Responses ranged from "I never use the numpad without looking at the keys" to "Very comfortable". Accuracy was measured as the proportion of entries that were typed correctly on a total of 30 data sheets that each contained 42 entries. A total of 98 undergraduate students participated in this study in exchange for course credit. Participants were randomly assigned to one of three data entry methods: single entry, single entry with visual checking, and double entry. The correlation between accuracy and comfort with the number pad was calculated for all participants combined and for participants within each of the three data entry conditions. None of these four correlations were statistically significant.

These non-significant results could be due to three factors. First, the accuracy data is extremely skewed, which limits the correlation it can have with other variables. Future research should use a more complicated data entry task, so that the accuracy data is less skewed. Second, it is possible that some participants lied about their comfort with the number pad. If so, this would add random variation to this variable, decreasing its correlation with other variables. Future research should use a different experimental procedure to encourage truthfulness: for example, if participants thought that the study administrator was not paying attention to their responses, they might answer more honestly. Finally, it may be that these results were non-significant because of the phrasing of the number pad question. Future research should assess both physical comfort with the number pad and how confident the participant is when using the number pad without looking.

INTRODUCTION

Inaccurate data entry introduces random error to a database. In laboratory settings, this reduces reliability, power, and effect sizes (Barchard & Pace, in press). In employment settings such as accounting and cashiering, inaccurate data entry could result in tax errors, poor management decisions, and incorrect payments (Healy, Kole, Buck-Gengler, & Bourne, 2000). Thus, accurate data entry is very important both within and outside of the research laboratory.

Many factors can influence individuals' data entry accuracy. These include speed, fatigue, repetition, and experience. Individuals who complete a data entry task with great speed are less accurate (Healy et al. 2000; Norr, 2000). Those who are subjected to extended periods of data entry become fatigued, which decreases accuracy (Leung, 2007). Similarly, when data entry is repetitious, errors increase (Healy et al. 2000). Finally, individuals with more computer experience are more accurate (Czaja, Shairt, Sanakaran, & Rubert, 1998).

Two studies have examined the effect of alternative physical arrangements on data entry performance. Woods and Baski-Reeves (2005) demonstrated that keyboard angle influences both physical comfort and typing speed and accuracy. Marteniuk, Ivens, and Brown (1996) examined four alternative number pad arrangements, but found no significant differences between them in terms of either speed or accuracy. No previous research has examined the relationship between comfort with the number pad and accuracy on a data entry tasks. We hypothesize individuals who are more comfortable with the number pad will be more accurate on the data entry task.

METHOD

Participants

A total of 98 (63 females, 35 males) undergraduate students participated in this study in return for course credit. Participants ages ranged from 18 to 44 (mean 20.78, standard deviation 4.78). In terms of ethnicity, 39.8% identified themselves as Caucasian, 17.3% as Asian, 15.3% as African American, 13.3% as Hispanic, 8.2% as Pacific Islander, and 4.1% as other.

Procedures

The study was administered on a computer during a 90-minute individual appointment. Participants were randomly assigned to one of three different data entry methods: Single Entry, Single Entry with Visual Checking, or Double Entry. Participants assigned to the Single Entry method were instructed to enter the data once. They were told that accuracy was more important than speed, and to please be as accurate as they could. Participants assigned to the Single Entry with Visual Checking method were asked to visually compare the paper data sheet to the values they had entered, after entering each sheet. They were asked to correct all errors that they found. Finally, participants assigned to the Double Entry method were instructed to enter each data sheet twice. The computer automatically compared the two data sets for mismatches and out-of –range errors. If an error occurred, the computer highlighted the cell that had the error, so that the participant could locate it easily. Participants were asked to check for mismatch and out-of-range errors, and correct all errors they found.

Participants entered data sheets which contain the following variables: ID number, Sex, and a sequence of ten entries for four different scales labeled "Family Background," "School Experiences," "Extraversion," and "Social Skills Test". Participants were asked to type numbers even though some of these variables contained letters. "Sex" was shown on the data sheets as "M" for male and "F" for female. Participants were asked to type "1" for male and "2" for female. In addition, on the "Family Background" scale, participants typed 1, 2, 3, 4, or 5, instead of the responses that were given on the data sheets, which read "SD", "D", "N", "A", and "SA". Finally, for the "School Experiences" scale, participants typed 1, 2, or 3, instead of "D", "N", or "A". The "Extraversion" and "Social Skills Test" scales were already in number form; thus participants were not asked to convert these scales.

Before beginning their data entry task, participants watched two video tutorials on the computer. The first video described the basics of using a Microsoft Excel™ spreadsheet. The second video taught the participant how to enter data according to their randomly assigned method. Next, participants completed data entry for five practice sheets. The study administrator monitored the participant, to ensure that they entered these data sheets correctly. For example, in the Single Entry with Visual Checking condition, if the participant did not visually check their errors, the administrator would remind the participant to do so, and in the Double Entry condition, the administrator would ensure that the participant corrected all of the highlighted errors. After the completion of the first set of data sheets, the participant saved and closed the spreadsheet and opened a new Excel spreadsheet for the main data entry task. During the main data entry task, participants entered thirty data sheets. The study administrator stayed in the same room, to be available to answer any questions the participant had through the remainder of the study. The participant then saved the file closed Excel. Finally, participants completed a short evaluation of the data entry method they used.

Measures

Comfort with the number pad was assessed using a self-report question that was asked at the beginning of the study: "How comfortable are you with using the number pad ("numpad") without looking at the keys?" Participants were given several possible ways to respond. If participants selected "I never use the numpad without looking at the keys" or "Very uncomfortable", this was coded as 1. "Moderately uncomfortable" was coded as 2, "Neutral" was coded 3, "Moderately comfortable" was coded 4, and "Very comfortable" was coded 5. If participants selected "Not applicable - my computer dos not have a numpad." then they were not included in this analysis.

Data entry accuracy was calculated as the percentage of entries that matched the data sheets. Participants entered 42 variables for each of the 30 data sheets, for a total of 1260 entries. For each participant, accuracy was calculated by the computer across all 1260 entries, for each participant.

RESULTS

When we calculated the correlation between accuracy and comfort with a number pad for all participants combined, the Pearson correlation was non-significant ($r(96) = -.03$, $p = .76$). When we calculated this correlation again for each different type of data entry, the three correlations were also non-significant (single entry $r(38) = .09$, $p = .64$; visual checking $r(39) = -.09$, $p = .58$; double entry $r(39) = -.30$, $p = .10$).

DISCUSSION

The purpose of this study was to determine if a relationship exists between accuracy on a data entry task and comfort with the number pad. In this study, there was no relationship between these two variables.

These non-significant results could be due to three factors. First, the accuracy data is extremely skewed: almost all participants were highly accurate. Most participants were 98% accurate or higher, and all but two participants were 94% accurate or higher. When a variable is highly skewed, this influences the size of the correlation it can have with other variables. Future research should use a more complex data entry task to reduce skew, and make it possible to find relationships between accuracy and other variables.

Second, it is possible that some participants lied about their comfort with the number pad, either to downplay their high skill levels or to exaggerate low skill levels. This is more likely in this study than in other studies because the research assistants who ran the study provided close supervision of participants during the first part of the study, when participants were answering this question. The study administrators needed to provide close supervision in order to determine which data entry method each participant had been assigned to, and in order to ensure that participants completed the data entry correctly during the practice session. However, this may have had the unintended side-effect of making participants lie about their comfort with the number pad. Future studies could assess comfort with the number pad (and other background questions) in a ways that increases privacy and thus encourages truthfulness. For example, if participants thought that the study administrator was not paying attention to their responses during this part of the study, they might answer more honestly.

Finally, it may be that these results were non-significant because of the phrasing of the number pad question. Participants were asked to indicate their comfort with the number pad. However, the word "comfort" may be interpreted as a physical condition or it could indicate how confident a participant is with using the number pad without looking. Future research could assess both physical comfort and confidence levels.

REFERENCES

Barchard, K.A. & Pace, L.A. (in press). Meeting the challenge of high quality data entry: A free double-entry system. *International Journal of Services and Standards*.

Czaja, S., Sharit, J., Sanakaran, N., & Rubert, M. (1998). Understanding sources of user variability in computer-based data entry performance. *Behaviour & Information Technology, 17,* 282-293.

Healy, A.F., Kole, J.A., Buck-Gengler, C.J., & Bourne, L.E. (2004). Effects of prolonged work on data entry speed and accuracy. *Journal of Experimental Psychology: Applied*, *10*, 188-199.

Leung, A.W. (2007). Neurophysiological correlates of performance and fatigue in study. *The Sciences and Engineering, 68*, 1349-1519.

Marteniuk, R., Ivens, C., & Brown, B. (1996). Are there task specific performance effects for differently configured numeric keypads? *Applied Ergonomics, 27,* 321-325.

Norr, K.B. (2000). An investigation of the relationship between speed and accuracy utilizing a data entry task (individual differences). *The Sciences and Engineering*, *60*, 3606-3772.

Woods, M. & Babski-Reeves, K. (2005). Effects of negatively sloped keyboard wedges on risk factors for upper extremity work related musculo-skeletal disorder and user performance. *Ergonomics, 48,* 1793-1808.