

Relationship between Computer Skills and Data Checking Accuracy

Jane C. Park, Brian W. Kautz, Kelly E. Grob, and Kimberly A. Barchard
University of Nevada, Las Vegas

Reference: Park, J. C., Kautz, B.W., Grob, K.E. & Barchard, K.A. (April, 2011). *Relationship between computer skills and data checking accuracy*. Poster presented at the Western Psychological Association Annual Convention, Los Angeles, CA.

Contact Information: Kim Barchard, Department of Psychology, University of Nevada, Las Vegas, 4505 S. Maryland Parkway, P.O. Box 455030, Las Vegas, NV, 89154-5030, USA, barchard@unlv.nevada.edu

Abstract

Errors in research data can change a study's results. Data checking is used to reduce the number of errors. Some people might be better than others when doing data checking. Previous research has shown that computer skills influence data accuracy, but has not examined which computer skills are the most important. The purpose of this study is to determine which computer skills have the strongest relationships with data checking accuracy. A total of 29 participants completed this study. First, they watched a five-minute video on how to use a spreadsheet. Then they checked a spreadsheet of data and corrected the errors they found. Accuracy was defined as the number of correct entries after the data checking was finished. Four computer skills were measured using self-report questions: typing without looking at the keys, using the number pad, using spreadsheets, and using computers. In this study, number pad skill had a strong significant correlation with data checking accuracy. Why was this skill the most important? All of our participants had some background skill in typing and using computers. Apparently, their level of knowledge was adequate. In addition, although some participants had more background with spreadsheets than others, the brief introduction to spreadsheets that we provided appears to have been adequate. Thus, the computer skill that had the strongest relationship with accuracy was one we did not teach, that takes a long time to acquire: number pad skill. Data checking accuracy is essential for valid scientific findings, and the people who do data checking should be the ones with the best number pad skill. Alternatively, people who will do data checking should improve their skill in this area.

Introduction

Data entry can introduce errors into otherwise sound data (Burchinal & Neebe 2006; Winkler, 2004). Errors can change a study's results and lead to invalid findings. Data checking is an effort to minimize these errors before the data is statistically analyzed (Burchinal & Neebe 2006). Because computers are used for data checking, computer skills may affect the quality of data checking. Previous research has shown that computer skills are correlated with data entry accuracy (Czaja, Sharit, Nair, & Rubert, 1998). This study will expand on previous research by including various computer skills, specifically, number pad skill, typing skill, spreadsheet skill, and computer skill.

This study examines the relationship between computer skills and data checking accuracy. This will be done by correlating the participants' computer skill and their performance during a data checking experiment. We hypothesize that all computer skills will have large positive correlations with data checking accuracy.

Method

Participants

A total of 29 participants (15 female and 14 male) completed the study for a course credit. The participants' ages ranged from 18 to 39 (mean 22.29 and SD 6.02). The participants identified themselves as African American 13.8%, Asian 27.6%, Caucasian 24.1%, Hispanic 24.1%, Pacific Islander 6.9%, and other 3.4%.

Procedures

Participants completed the measures on the computer during a 90-minute supervised testing session. Participants watched training videos for Microsoft Excel and for one of three randomly assigned data checking methods: visual checking, read aloud, and double entry. In all three conditions, participants tried to identify the errors that exist in a Microsoft Excel worksheet. In visual checking, participants were given a paper copy of the original data and compared it visually with the data on the Excel worksheet. If they noticed any mismatches, they were asked to correct the Excel worksheet. In read aloud, the participant looked at the data on the Excel worksheet while the researcher read the paper sheet aloud. When the data read by the researcher differed from the data in the Excel worksheet, the participant said "verify," to prompt the researcher to read the data again. If the Excel sheet disagreed with the paper data sheet, the participant fixed the error. Finally, in double entry, participants entered the data a second time. If the data entered by the participant did not match the data that was already given in the Excel worksheet, Excel highlighted the differences. Excel also highlighted any entries that were outside the allowable range for each variable. Participants were instructed to correct the errors that were identified.

Measures

The participants checked 20 data sheets with a total of 680 entries. Data checking accuracy was calculated as the number of correct entries after the participant had finished data checking. Correct entries were defined as an entry in the Excel sheet that corresponded with the original paper sheet.

Computer skill was measured by asking the participant to rate their ability on a five point scale (Very Low, Low, Moderate, High, and Expert). They were asked how they would gauge their ability to “Use computers,” “Use spreadsheet programs (e.g., Excel, Works Spreadsheet),” Type without looking at the keys,” and “Use the numeric keypad (numpad) without looking at the keys.”

Data Analysis

To examine the relationship between computer skill and accuracy, we correlated the number of correct entries with the four computer skills.

Results

All computer skills had small to moderate positive correlations with data checking accuracy. However, only number pad skill had a significant correlation. See Table 1.

Table 1

Correlations with Number of Correct Entries

Ability	Correlation
Computers Ability	.18
Spreadsheet Ability	.24
Type Ability	.39
Number Pad Ability	.42*

* $p < .05$.

Discussion

The purpose of this study was to determine whether computer skills affect data checking accuracy. We believed that all of the computer skills would have high correlations with data checking accuracy. However, only number pad skill was had a statistically significant correlation, and this correlation was only moderate. Why was it that only this skill had a significant relationship with accuracy? Using the number pad is a skill that was not touched upon during training, but appears to be an area upon which there are important differences. In contrast, all participants had a satisfactory amount of general computer knowledge and general typing skill before coming into our study, and any differences in spreadsheet skills was nullified by the training given at the beginning of the study.

An implication of our findings is that they show us which individuals are more statistically prone to make data checking errors. We can use a person’s self-assessed computer skills to determine if they need additional training or need to be barred from data checking in order to reduce errors.

Future research should examine the relationship between computer skills and each of the data checking methods, Read Aloud, Visual Checking, and Double Data Entry. This will indicate if computer skills affect the individual methods in different ways. The results may show which method has stronger correlation with each of the computer skills. And with these results, training can be geared towards specific data checking methods. The person performing data checking can be assigned to one of the data checking methods depending on their computer skills. For example, if there was a correlation between Double Data Entry and typing skills, someone who scored high on typing skills should be assigned Double Data Entry.

References

- Burchinal, M., & Neebe, E. (2006). I. Data management: Recommended practices. *Monographs of the Society for Research in Child Development*, 71(3), 9-23. doi:10.1111/j.1540-5834.2006.00354.x
- Czaja, S., Sharit, J., Nair, S., & Rubert, M. (1998). Understanding sources of user variability in computer-based data entry performance. *Behaviour & Information Technology*, 17(5), 282-293. doi:10.1080/014492998119355
- Winkler, W. E., (2004). Methods for evaluating and creating data quality. *Information Systems*, 29, 531–550.