Researchers use multiple methods for data checking. Each method can help identify and fix errors that were introduced during the data entry process. Fixing the errors that were introduced during the data entry process increases the accuracy of the research results. Accuracy is important because if a researcher publishes inaccurate results other researchers would not be able to replicate those results and draw the same conclusions. The purpose of this study is to compare the accuracy of four different data checking methods: double entry with one person, double entry with two people, visual checking, and solo read aloud. So far, previous research has shown that double entry is more accurate than visual checking (Barchard & Pace, 2011) and partner read aloud (Kawado, Hinouto, Matsuyama, Yamauchi, Hashimoto, & Ohashi, 2003). Although there has not been many studies done on the comparison of these four methods and only one study has used solo read aloud, double entry has been shown to produce the highest quality data. I therefore hypothesize that the two double entry methods will have the highest accuracy.

There are four common data checking methods: single person double entry, two person double entry, read aloud, and visual checking. The single person double entry method consists of the one person entering and checking data. The two person double entry method has one person entering the data and a second person entering the data a second time and checking that they match. The read aloud method has one person entering the data and either the same person (solo read aloud) or a different person (partner read aloud) checking the data by reading it aloud. The visual checking method consists of one person entering the data and checking the data visually. One study that compared three different data checking methods found that two person double entry produces fewer errors but takes longer than other data checking methods (Barchard & Verenikina, 2013). Through comparing the accuracy of the four different data checking methods this study will be able to identify which method produces the fewest errors.

Participants

errors will be 100 participants for each data checking method, giving a total of 400 participants. Participants will be undergraduates at the University of Nevada, Las Vegas. Participants will be recruited from the Department of Psychology subject pool.

Materials

Measures

Accuracy will be measured by the numbers of errors on the participant’s final Excel sheet. An error is defined as a discrepancy between the Excel sheet and what was actually on the data sheets. The data checking method that produces the greatest number of errors will be considered the least accurate. The method that produces the least number of errors will be considered the most accurate data checking method.

Data Analysis

To compare the accuracy of the four data checking methods, an ANOVA will be calculated. The independent variable will be the group each participant belongs to (one-person double entry, two-person double entry, solo read aloud, or visual checking). The dependent variable will be the number of errors left in the Excel sheet after the participant has completed entering and checking data.

Discussion

Unlike previous studies, this study will compare four data checking methods simultaneously. One method that this study includes is solo read aloud, for which there has been very little published research. Only a single study has examined solo read aloud and in that study, it was only compared to only one other data checking method: double entry (Kawado et al., 2003). Moreover, that study used only two participants. In our study, we will be comparing 100 participants in solo read aloud to 300 participants in the other three data checking methods.

One weakness of our study is that we are not including every possible data checking method: We are excluding partner read aloud. Partner read aloud is a data checking method similar to solo read aloud. In partner read aloud, there are two people checking the data. One person reads from the original data sheet while the other person visually checks the data in the Excel file. Partner read aloud has been excluded in order to simplify administration procedures and reduce the time it will take to complete the study. However, because we are excluding partner read aloud, we are not comparing all data checking methods that are available to researchers. A better study would be to compare the accuracy of all possible data checking methods.

We are excluding partner read aloud. Partner read aloud is a data checking method similar to solo read aloud. In partner read aloud, there are two people checking the data. One person reads from the original data sheet while the other person visually checks the data in the Excel file. Partner read aloud has been excluded in order to simplify administration procedures and reduce the time it will take to complete the study. However, because we are excluding partner read aloud, we are not comparing all data checking methods that are available to researchers. A better study would be to compare the accuracy of all possible data checking methods.

I would like to thank my mentor Dr. Kimberly Barchard for guiding me through the process of research and allowing me to comprehend and understand our research topic to the fullest. Also, I would like to thank the McNair Program for providing workshops and guidance through the McNair summer research program. Lastly, I would like to thank my peers who worked alongside me in the Interactive Measurement Group.

References

Paulsen, A., Overgaard, S., & Lauritsen, J. (2012). Quality of data entry using single entry, double entry and automated forms processing—an example based on a study of patient- registered interventions. PLoS ONE, 7(10), e45023. doi:10.1371/journal.pone.0045023