Abstract

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, Hinotsu, Matsuyama, Yamaguchi, 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.

Introduction

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.

Discussion

Unlike previous studies, this study will compare four data checking methods simultaneously. One method that this study will not use is the solo read aloud method. The only previous study that compared solo read aloud to another method was only compared to only one other data checking method: double entry (Kawado et al., 2003). Moreover, that study only used two participants. In our study, we will be comparing 100 participants in both the solo read aloud and the double entry method. 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 while the other person visually checks the data in the Excel file. Partner read aloud has been used in order to simplify the 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.

Methods

Participants

There 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

There have been many different data checking methods. These include solo read aloud, partner read aloud, visual checking, and double entry with one or two people. Double entry has consistently been shown to be the most accurate (Barchard & Pace, 2011; Gibson, Harvey, Everett, & Parmar, 1994; Paulsen, Overgaard, & Lauritsen, 2012) and has been defined as the definitive gold standard of good clinical practice (Paulsen et al., 2012). However, even though double entry may find the most errors, not all researchers have concluded that it benefited their research (Gibson et al., 2012). The biggest problem with the double entry method is that it takes up a greater amount of time. Double entry requires up to 37% more time than other data checking methods (Reynolds-Haertle & McBride, 1992). Regardless of which data checking method is used, accuracy rates increase when the data checking person is someone different from the data entry person. One study found that read aloud detected about 60% of the errors when a different person did the checking, but only 39.9% of errors when the same person did the checking and the original entering. Similarly, double entry detected 88.3% of errors using different operators, but only 69% of the errors when the data checking person was the same as the original data entry person (Kawado et al., 2003). Because of this, we hypothesize that double entry will be more accurate than read aloud and visual checking, and that double entry with two people will be more accurate than double entry with one person in the present study.

Measure

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.

Literature Review

There have been many different data checking methods. These include solo read aloud, partner read aloud, visual checking, and double entry with one or two people. Double entry has consistently been shown to be the most accurate (Barchard & Pace, 2011; Gibson, Harvey, Everett, & Parmar, 1994; Paulsen, Overgaard, & Lauritsen, 2012) and has been defined as the definitive gold standard of good clinical practice (Paulsen et al., 2012). However, even though double entry may find the most errors, not all researchers have concluded that it benefited their research (Gibson et al., 2012). The biggest problem with the double entry method is that it takes up a greater amount of time. Double entry requires up to 37% more time than other data checking methods (Reynolds-Haertle & McBride, 1992). Regardless of which data checking method is used, accuracy rates increase when the data checking person is someone different from the data entry person. One study found that read aloud detected about 60% of the errors when a different person did the checking, but only 39.9% of errors when the same person did the checking and the original entering. Similarly, double entry detected 88.3% of errors using different operators, but only 69% of the errors when the data checking person was the same as the original data entry person (Kawado et al., 2003). Because of this, we hypothesize that double entry will be more accurate than read aloud and visual checking, and that double entry with two people will be more accurate than double entry with one person in the present study.