Visual Checking: Why Take The Risk?

Daniel N. Erosa, Ashley A. Anderson, Yevgeniya Verenikina, & Kimberly A. Barchard
University of Nevada, Las Vegas

Reference: Erosa, D.N., Anderson, A.A., Verenikina, J., & Barchard, K.A. (April, 2012). *Visual Checking: Why Take The Risk?* Poster to be presented at the Association for Psychological Science, Chicago, IL.

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

Double entry is more accurate than other data checking methods. Is the greater accuracy due to the increased time spent on data checking? 188 participants were randomly assigned to three data checking methods. After controlling for time, double entry still corrected more errors than read aloud or visual checking.

Introduction

When entering data, there are several errors that can arise. In order to account for these errors, researchers use various methods of data checking to ensure accuracy. The method of data checking used is important because accurate data checking is critical to having accurate results. For example, a single error that has been over looked by the data checking process can turn a moderate correlation to zero, or make a significant *t*-test non-significant (Barchard & Pace, 2011). That being said, there are several different methods of data checking to choose from.

Three commonly used methods of data checking are: read aloud, visual checking, and double entry. During the read aloud method, one person reads the original data out loud while someone else checks the entered data. The visual checking method involves someone checking data that has already been entered to the original data. Finally, the double entry method involves someone entering the data twice while the computer compares the two entries for mismatches or values outside the range.

Researchers may prefer to use visual checking because it is faster than double entry (Kawado et al., 2003), but that does not mean it is more accurate. In fact, visual checking is the least accurate of the three data checking techniques (cite) . Research shows that double entry takes longer than other data checking methods, but is more accurate (Reynolds-Haertle & McBride 1992). . Does this mean that double entry is more effective due to the fact that it takes longer? Previous research has not yet explored this hypothesis. The purpose of this study is to examine if the reason that double entry is so accurate is due to the time that people spend using double entry.

Method

Participants

A total of 117 participants (? males and ? females) participated in this study for credit towards their psychology course. The age of the participants ranged from 18 to 67 (mean 21.8, standard deviation 6). Participants included Caucasian (42.6%), Asian (21.8%), Hispanic (15.4%), African American (12.2%), Pacific Islander (5.3), and Other (2.1%)

Procedures

Participants completed a 90 minute session while being supervised by a researcher. The participants watched a short instructional video on how to use Microsoft Excel. Then, the participants were randomly assigned one of three data checking methods. The participants watched a second instructional video on how to use their assigned data checking method. The three methods that were used are: visual checking, double entry, and read aloud. After the video, the participants started part one of the study. The participants practiced their assigned techniques by checking five data sheets and correcting any errors. The data sheets contained only letters and numbers. During part two of the study, 20 additional sheets of data were presented to the participants to check. At the end of the study, participants were asked to answer a short questionnaire. The questionnaire contained questions involving: demographic information, computer skills, and experience during the study, and their perception of the method they were using.

Measures

Accuracy was measured by the number of correct entries in part 2 of the study. An entry was considered correct if it was identical to the original data. Time was measured by recording the time the participant started part 2 of the study until the end of part 2.

Analysis

To test for sex differences on the Metaphors Test, we used an independent samples t-test. The grouping variable was sex and the dependent variable was the total scores on the Metaphors Test.

Results

Visual checking was faster than read aloud, which was faster than double entry. We compared the time it took to complete the data checking using the three methods using a one-way ANOVA. The overall ANOVA was significant, $F(2, 114) = 22.46$, $p < .001$. Tukey's HSD showed that each of the methods was significantly different from the other two. See Table 1 for the mean time for each of the three methods.

Visual checking was less accurate than read aloud, which was less accurate than doube entry. We compared the three methods in terms of the accuracy, using a one-way ANOVA. The overall ANOVA was not quite significant, $F(2, 114) = 2.41$, $p = .094$. However, the mean accuracy values showed the expected differences, with double entry being the most accurate and visual checking being the least accurate. See Table 1 for these three means.

Finally, we wanted to determine if the greater accuracy of the double entry method could be explained by the additional time spent on the data checking. A two step process was used. First, we predicted accuracy using time. We conducted this analysis to get the residuals; which is a pure measurement of accuracy once time has been taken into account. Second, we compared the residuals across the data checking conditions using one-way ANOVA. The results were significant, $F(2, 185)= 40.99$, $p < .001$. Tukey's HSD showed that all pair-wise differences were significant at p<.001. Once we have controlled for time, visual checking was still less accurate than read aloud, which was still less accurate than double entry. See Table 1 for the means of each group.

Table 1

Comparing Three Data Checking Methods

Method	Time (in min)	Raw Accuracy	Accuracy Controlling for Time
Double Entry	39.49	679.71	678.02
Read Aloud	34.78	677.69	677.71
Visual Checking	28.00	675.11	677.28

Discussion

Previous research has shown that visual checking is less accurate than read aloud, which is less accurate than double entry. We replicated those findings, showing that visual checking is faster but less accurate than read aloud, which is faster but less accurate than double entry. The purpose of this study was to see if the time spent on each method of data checking can account for the differences in accuracy. Our hierarchical multiple regression show that the differences in accuracy remained significant after controlling for time. Tukey's HSD showed that the differences between the conditions is significant, regardless of which condition was paired with the other, and that double entry was still the most accurate and visual checking still being the least accurate. These results support to the notion that double entry is not only the most accurate data checking technique but also the most valuable with one's time.

The results of this test could persuade future researchers to invest the extra time it takes to use double entry. The risks of using a less accurate method, such as visual checking or read aloud, can have serious implications of the data being analyzed. These implications can cause significant results to be interpreted as insignificant results due to incorrect data entry. The results of this study may also be of interest for those businesses that have a paid position for data entry. Using the most effective method of data checking, double entry, can save a business a lot of time, money, any worry of erroneous entries.

Future research should examine any other possible data checking methods that have not yet been studied, such as solo read aloud. This may provide assist in finding the best data checking technique.

Future research could also examine the attention span of the participant too see if it has a positive or negative correlation with the accuracy of the data checking method. If differing attention spans are found to have a significant effect on accuracy, it may change the way the studies are conducted. These findings could also provide support that double entry is the best tool for data checking since the program compensates for human overlook or human error.

References

Barchard, K. A. & Pace, L. A. (2011). Preventing human error: The impact of data entry methods on data accuracy and statistical results. *Computers in Human Behavior, 27*(5), 1834-1839. doi:10.1016/j.chb.2011.04.004.

Kawado, M., Hinotsu, S., Matsuyama, Y., Hashimoto, S., & Ohashi, Y. (2003). A comparison of error detection rates between the reading aloud method and the double data entry method. *Controlled Clinical Trials*, *24*(5), 560-569. doi: 10.1016/S0197-2456(03)00089-8

Reynolds-Haertle, R. A., & McBride, R. (1992). Single versus double data entry in CAST *Controlled Clinical Trials*, *13*, 487-494.