To Err is Human:
To Catch Those Errors is Good Research Practice

Tina Tran, Misha B. Asif, Brianna R. Maxim,
and Kimberly A. Barchard

Reference: Tran, T., Asif, M. B., Maxim, B. R., and Barchard, K. A. (2015, April). To err is human: To catch those errors is good research practice. Paper presented at UNLV Psi Chi Symposium, Las Vegas, NV.

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

Abstract

Errors that occur during data entry can have profound implications on statistical results and study conclusions. Therefore, researchers spend valuable time checking their data. This paper compared four data checking methods: double entry, visual checking, solo read aloud, and partner read aloud. A total of 154 undergraduates (103 female) participated in this study. Participants were randomly assigned to one of the four data checking methods. During the 90-minute session, the participant checked 20 data sheets, each of which contained 26 data points. These data sheets had already been entered into Excel, but we deliberately introduced errors into the Excel sheet; the participant's job was to find and correct these errors. Partner read aloud had significantly more errors than the other methods. Double entry had the fewest errors (although the differences with solo read aloud and visual checking were not significant). Future research should replicate this study in non-undergraduate populations.

Introduction

Data checking remains an important aspect of research that should not be overlooked. The quality of data collection, as well as statistical analyses, are dependent on having adequate data checking techniques. If the integrity of the data is questionable, it will be reflected in the results (i.e., obtaining low quality data filled with high level of errors can lead to misinterpreted results as well as a waste of resources; Moody & McMillian, 2002). Schneider and Deenan (2004) emphasized the need for reducing quantitative data errors in order to avoid compromising research findings. Data entry errors are costly and take away from the reliability and validity of the research (Atkinson, 2012; Day, Fayers, & Harvey, 1998). For example, Goldberg, Niemierko, and Turchin (2008) found that when they analyzed several clinical nursing databases, many errors came from manual data entry mistakes as well as misinterpretations of data that originated from hard copies. These types of errors had implications about the quality of patient care (Goldberg et al., 2008). Similarly, Burchinal and Neebe (2006) found that out of the 5 tests of 80 children, 55% of test scores were erroneous and 20% had serious mistakes. The purpose of this study is to compare four data checking methods: Double entry, visual checking, solo read aloud, and partner read aloud.

Three studies have shown that double entry results in fewer errors than the other techniques. First, Kawado, Hinotsu, Matsuyama, Yamaguchi, Hashimoto, and Ohashi (2003) compared double entry and partner read aloud. They found that double entry detected 88.3% of errors, while partner read aloud detected only 59.5% errors. They concluded that double entry was superior when compared to partner read aloud. Second, Barchard and Pace (2011) concluded that double entry resulted in significantly fewer errors when compared to visual checking and single entry. Visual checking resulted in 2958% more errors than double entry. Finally,

Barchard and Verenikina (2013) examined double entry, solo read aloud, and visual checking, and concluded that double entry was superior to read aloud and visual checking among people with and without previous data entry experience. Read aloud and visual checking contained more than 20 times as many errors compared to double entry (Barchard & Verenikina, 2013).

We therefore hypothesize that double entry will have the fewest number of errors in our study.

Method

Participants

A total of 154 participants took part in this study (49 male, 103 female, 2 unspecified). Participants were undergraduate students, who earned 1.5 research credits for their participation. The ages ranged from 18-50 years old (mean 21.56, SD 6.42). These individuals identified their ethnicities as 10.4% African American, 20.1% Asian, 36.4% Caucasian, 26% Hispanic, 0.6% Native American, 1.9% Pacific Islander, and 3.9% other.

Procedures

Before participants arrived, the data from 25 data sheets were entered into Excel. We deliberately introduced errors into the Excel file. The participants' task was to locate and correct any errors they found.

The data sheets (see Figure 1) contained 26 pieces of data. Some of these data points exactly matched what had been entered into Excel. Others required conversion. Specifically, sex was entered as "1" for M and "2" for F, and the five options for the Study Habits section were entered as 1, 2, 3, 4, and 5. These formats are all common in psychological studies. Conversions were used because errors are more common (both typing errors and translation errors are possible), thus creating a more stringent test of the data checking methods.

Participants were randomly assigned to one of four data checking methods: double entry, visual checking, solo read aloud, and partner read aloud. After watching a video that provided an overview of Excel, they watched a second video that explained their data checking method. To practice using their data checking method, they checked the data from five data sheets. See Figure 1. In the main part of the

The Learning Study

ID: 739925

Sex: (M) F

	Learning Style		Study Habits	
1.	1 2 3 (4) 5	1.	SD D N A (SA)
2.	1 2 3 4 (5)	2.	SD D (N) A SA
3.	1 2 3 4 (5)	3.	SD D N A (SA)
4.	1 2 (3) 4 5	4.	SD D N A (SA)
5.	1 2 3 4 (5)	5.	SD D N (A) SA
6.	1 2 3 (4) 5	6.	SD D N (A) SA
7.	1 2 3 4 (5)	7.	SD (D) N A SA
8.	1 2 (3) 4 5	8.	SD (D) N A SA

	Spelling Test		Math Test
1.	ACCOMMODATE	1.	156
2.	AMATEUR	2.	235
3.	CALENDAR	3.	485
4.	CEMETERY	4.	493
5.	CONSHENCE	5.	364
6.	EMBARRASS	6.	327
7.	EXHILARATE	7.	203
8.	MAINTAINANCE	8.	347

Figure 1: Example of Data Sheet

study, they checked an additional 20 sheets. Sessions lasted about 90 minutes.

In the double entry condition, participants entered the data a second time. After each row was completed, Excel compared the data the participant had entered with the original data. Discrepancies were highlighted. Participants corrected any errors that they noticed.

In the visual checking condition, participants visually compared the data presented on the computer screen with the data on the paper data sheet. Again, if there were any inconsistencies, the participant corrected the errors.

In the solo read aloud condition, participants read the data sheets out loud. Then they looked at the Excel file to see if it matched. If they noticed discrepancies, they correct them.

In contrast, in the partner read aloud condition, the study administrator read the data sheets out loud, while the participant looked at the Excel file to see if they matched. If the participant heard what appeared to be discrepancy, they asked the administrator to read that data point again. If they confirmed there was a discrepancy, they corrected the error.

Data Analysis

To examine the quality of the four data checking methods, we compared the number of errors in the Excel sheets after data checking had been completed. These errors could have been in the Excel sheet when the participants started data checking, and they simply failed to correct them. Alternatively, these errors may have been introduced by the participants themselves.

To compare the number of errors across the four data checking methods, we used a one-way analysis of variances. When there was a significant difference between the means, Tukey post-hoc tests were also conducted to determine which specific means differed.

Results

There were differences in the average number of errors left behind among the four methods ($F(3, 150) = 6.39$, $p < .001$). See Table 1. Tukey's post-hoc tests showed that partner read aloud resulted in more errors than double entry ($p < .001$), visual checking ($p = .20$), and solo read aloud ($p = .047$). No other pairwise differences were statistically significant, although we noted that double entry had far fewer errors than the other methods. On average, compared to double entry, visual checking had 3.06 times as many errors, solo read aloud had 3.38 times as many errors, and partner read aloud had 7.51 times as many errors.

Table 1
Number of Errors in the Data Set after Checking

Method	Mean	Standard Deviation
Double Entry	0.47	1.59
Visual Checking	1.44	2.22
Solo Read Aloud	1.59	2.19
Partner Read Aloud	3.53	5.15

Discussion

The purpose of this paper was to determine which data checking method removes the most errors. We examined four data checking methods: double entry, visual checking, solo read aloud, and partner read aloud. We hypothesized that double entry would be the most accurate based on previous papers (Barchard & Pace, 2011; Barchard & Verenikina, 2013; Kawado et al., 2003; Reynolds-Haertle, & McBride, 1992). The results did not support this hypothesis. Although double entry had fewer errors than the other methods, the differences were not statistically significant. The only significant difference was that partner read aloud left significantly more errors in the data compared to the other methods.

The implications of this finding are clear: Do not use partner read aloud. It resulted in more than twice as many errors as the next worst technique (solo read aloud) and it takes roughly twice as many resources, because it required two research assistants to be involved.

Future research should replicate this study using participants other than undergraduate students. Many of our participants were freshmen (44.2%) and sophomores (27.3%), a population usually inexperienced in data entry and data checking. Future research should study professionals and graduate students who have already been trained in data collection, entry, and checking. In addition, future studies should use larger sample sizes. It is possible that the lack of significant differences between double entry and the other techniques were due to a lack of power.

References

Atkinson, I. (2012). Accuracy of data transfer: Double data entry and estimating levels of error. *Journal of Clinical Nursing, 21*(19-20), 2730-2735. doi:10.111/j.1365-2702.2012.04353.x

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

Barchard, K. A., & Verenikina, Y. (2013). Improving data accuracy: Selecting the best data checking technique. *Computers in Human Behavior, 29*(5), 1917-1922. doi:10.1016/j.chb.2013.02.021

Burchinal, M., & Neebe, E. (2006). Data management: Recommended practices. *Monographs of the Society for Research in Child Development, 71*(3), 9-23.2006-23007-003. doi:10.1111/j.1540-5834.2006.00402.x

Day, S., Fayers, P., & Harvey, D. (1998). Double data entry: What value, what price? *Controlled Clinical Trials, 19*(1), 15-24. doi:10.1016/s0197-2456(97)00096-2

Goldberg, S. I., Niemierko, A, & Turchin, A. (2008). Analysis of data errors in clinical research databases. *AMIA Annual Symposium Proceedings Archive 2008*, 242-246.

Kawado, M., Hinotsu, S., Matsuyama, Y., Yamaguchi, T., 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

Kraenzle Schneider, J., & Deenan, A. (2004). Reducing quantitative data errors: tips for clinical researchers. *Applied Nursing Research, 17*, 125-129. doi:10.1016/j.apnr.2004.02.001

Moody, L.E., & McMillan, S. (2002). Maintaining data integrity in randomized clinical trials. *Nursing Research, 51*(2), 129-133. doi:10.1097/00006199-200203000-00010

Reynolds-Haertle, R., & McBride, R. (1991). Single vs. double data entry in CAST. *Controlled Clinical Trials*, 487-494. doi:10.1016/0197-2456(92)90205-e