Double Entry and Read Aloud Identify More Errors than Visual Checking

Heather C. Johnson, Yevgeniya Verenikina, Spencer Hensley, Kelly E. Grob, and Kimberly A. Barchard
University of Nevada, Las Vegas

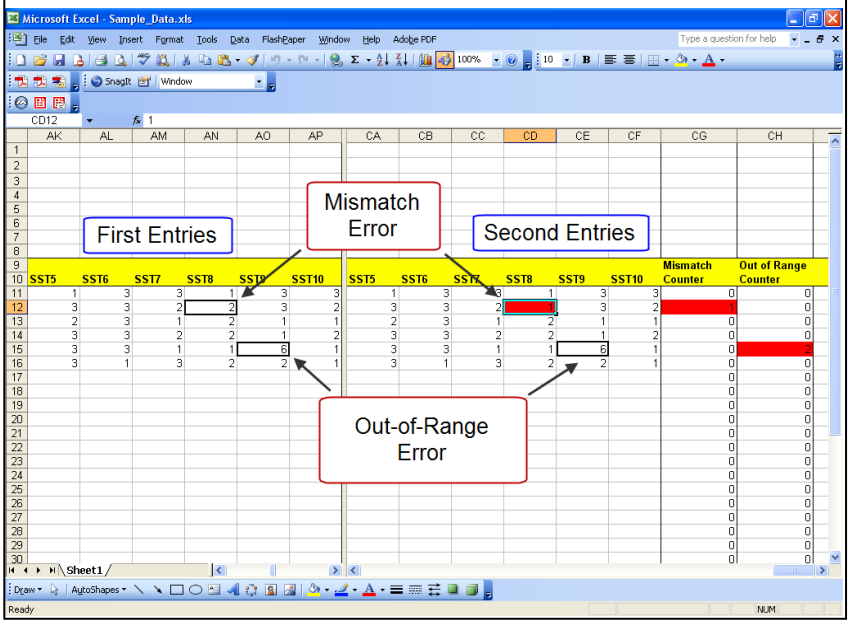
Reference: Johnson, H. C., Verenikina, Y., Hensley, S., Grob, K. E., & Barchard, K. A. (2011, May). *Double entry and read aloud identify more errors than visual checking.* Poster presented at the May 2011 Convention of the Association for Psychological Science, Washington, DC.

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

Accurate data are essential for valid science. Participants were randomly assigned to three data checking techniques. Double Entry and Read Aloud identified and eliminated significantly more errors than Visual Checking. Visual Checking resulted in 14 times as many errors left in the data set. Researchers should stop using Visual Checking.

Figure 1
Double Entry Screen Layout



Introduction

The validity of our research conclusions depends upon the accuracy of our original data. When data are typed into the computer, errors can be introduced. Small data entry errors can result in major changes in our conclusions. In the professional world, Winkler (2004) shows that small errors can have large monetary consequences. In the scientific arena, Burchinal and Neebe (2006) note that data entry errors reduce statistical power. High-quality data management practices are needed to maintain data integrity (Burchinal & Neebe, 2006).

The purpose of this paper is to examine the efficacy of three common data checking techniques: Double Entry, Visual Checking, and Read Aloud. In the Double Entry technique, users are instructed to enter the data a second time. The computer indicates if the two sets of entries match, and if not, the user corrects the errors. See Figure 1. In the Visual Checking technique, the entries are visually compared with the original paper data sheets, and errors are corrected. Finally, in the Read Aloud technique, the data is read aloud by either a person or a computer, while another person looks at the entries on the computer. Once again, when errors are noticed, they are corrected.

Previous research has shown that Double Entry leads to fewer errors than Read Aloud when paid professionals enter medical data (Kawado, Hinotsu, Matsuyama, Yamaguchi, Hashimoto, & Ohashi, 2003). However, no research has compared Double Entry to Read Aloud when research data are entered by data entry personnel who are unpaid volunteers (like the research assistants who usually enter academic research data). Moreover, no research has compared either technique to Visual Checking. The current paper seeks to fill this gap. The purpose of the current study is to compare these three techniques, using unpaid volunteers who are entering research data.

Method

Participants

Thirty-eight (20 female, 18 male) undergraduate students participated in this study in exchange for course credit. Ages ranged from 18 to 39 (mean = 22.27, SD = 5.58). Participants identified themselves as follows: Caucasian 34%, Asian 29%, Hispanic 18%, African-American 11%, Pacific Islander 5%, and Other 3%.

Equipment and Materials

The paper data sheets each contained 34 pieces of data. See Figure 2 for an example data sheet. Before participants arrived, these data were entered into a Microsoft Excel 2007 worksheet. However, the researchers deliberately introduced discrepancies between the paper data sheets and the Excel entries. The participants' task was to locate and correct these errors. Participants used a standard, non-ergonomic keyboard with a separate number pad on the right-hand side and a 17-inch color CRT monitor. The instructional videos were viewed using Flash 10 and Internet Explorer 8 and included sound.

Figure 2
Example Data Sheet

The Learning Study

ID: 739925

Sex: (M) F

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

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

Procedure

Participants completed the study during an individually administered, in person, 90-minute session, which was supervised by a trained administrator. Participants began by viewing an instructional video about how to use Excel. Participants were then randomly assigned to one of the three data checking techniques, and were shown a second instructional video explaining the particular technique to which they had been assigned. In the Double Entry technique, participants entered the data a second time, and the computer then compared these entries with the previous entries. If the entries were different, the computer highlighted the error. In the Read Aloud technique, participants read the previously entered data on the computer screen while the study administrator read the original paper responses out loud. If participants noticed an error, they were asked to say "verify" to prompt the administrator to read that data point again. In the Visual Checking technique, participants visually compared the data on the computer screen with the original paper responses. In all three conditions, participants were asked to correct all the errors that they noticed.

Participants began with five practice data sheets. While participants entered these sheets, the administrator offered to answer questions. These sheets were not scored. After completing these practice sheets, the participants then checked 20 additional data sheets. Finally, participants answered two brief questionnaires.

Measures

The dependent measure was the number of correct entries remaining after a participant finished the study. A *correct entry* was defined as an Excel entry whose contents match the original paper data sheet. To calculate the number of correct entries, participants' Excel files were imported into SPSS for scoring.

Data Analysis

Originally, we planned to use a one-way ANOVA to compare the accuracy of the three groups. However, severe heterogeneity of variance made this impossible. We therefore used a non-parametric statistic instead, the Kruskal-Wallis H Test. We compared the mean ranks of good entries for the three techniques.

Results

Double Entry and Read Aloud resulted in nearly perfect datasets. The mean number of correct entries was 679.53 and 678.86, respectively, out of a total of 680. In contrast, the Visual Checking technique had relatively low accuracy rates. The mean number of correct entries was 667.58. Stated another way, Visual Checking resulted in 25 times more errors being left in the dataset than Double Entry and 10 times more errors than Read Aloud. See Figure 3. These accuracy rates were significantly different ($H(2) = 14.07$, $p = .001$). See Table 1. A pairwise comparison of the techniques showed that Double Entry was significantly more accurate than Visual Checking. See Table 2.

Figure 3
Average Number of Errors for Each Data Checking Technique

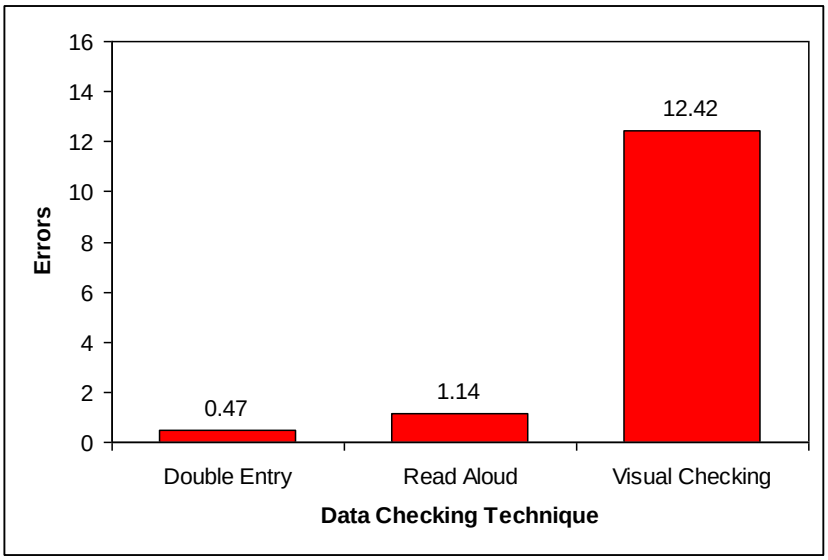


Table 1
Mean Rank of Correct Entries by Technique Checking Technique

Technique	N	Mean Rank
Double Entry	19	25.05
Read Aloud	7	19.50
Visual Checking	12	10.71
Total	38	

Table 2
Pairwise Comparison of Techniques

Technique Pairing	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj. Sig.
Visual Checking & Read Aloud	8.79	4.93	1.78	.075	.224
Visual Checking & Double Entry	14.34	3.82	3.75	.000	.001
Read Aloud & Double Entry	5.55	4.59	1.21	.226	.678

Discussion

Double Entry and Read Aloud are more effective than Visual Checking. In this study, we found that Visual Checking resulted in 25 times more errors being left in the dataset than Double Entry, and 10 times more errors than Read Aloud. The difference between Visual Checking and Double Entry was statistically significant, but the difference between Read Aloud and Double Entry was not. This is in contrast to previous research, which found that Double Entry was more accurate than Read Aloud (Kawado et al., 2003). However, data collection is on-going and we expect this difference to be significant by the time we reach our final sample size. We therefore recommend researchers use Double Entry or Read Aloud, and avoid Visual Checking.

Why is Visual Checking less effective than Double Entry and Read Aloud? Visual Checking relies on the participants to detect errors themselves, while the Read Aloud and Double Entry techniques assist the participants in identifying errors. In the Read Aloud technique, the administrator reads the data to the participant, one data point at a time. When participants notice discrepancies, they ask the researcher to read that data point again. Thus, the two people collaborate to identify errors. In the Double Checking technique, participants collaborate with the computer. The computer calls participants' attention to mismatched cells in Excel. Collaboration has been shown to increase the probability of detecting errors (Nihei, Terashima, Suzuki,

& Morikawa, 2002). In the Visual Checking technique, participants have no help in finding the errors. Relying solely on the user does not work. Researchers should stop using Visual Checking.

This study did not find a significant difference between the Double Entry and Read Aloud. Both had very high accuracy rates. Future research should try to differentiate these methods. First, there may be differences in time. Double Entry takes one person a bit longer, but the Read Aloud technique we used requires two people. Read Aloud could be done by having speech synthesis software read the data, which would make it faster than Double Entry. When paid professionals entered medical data, Double Entry was more accurate than Read Aloud with speech synthesis (Kawado et al., 2003) – but it might be that these two techniques are equally effective in a research context. Read Aloud could also be done by a single person by combining it with Visual Checking – a single person could read the data point on the paper data sheet out loud, and then visually check it against the entry on the computer screen, perhaps reading that entry out loud too. No research has examined the effectiveness of this technique. Second, there may be differences between Double Entry and Read Aloud when they are used in distracting environments. In this study, data checking was completed in a quite room, containing just the participant and the experimenter. In both business and research environments, data checking might often occur in much noisier environments. Future research could mimic those environments by playing background music or by conducting the data checking in a busy computer lab.

Finally, even if differences between Double Entry and Read Aloud do not generalize from the medical context to the research context, there could be other reasons for preferring one method over the other. For example, there could be differences in subjective evaluations of these two techniques. If the two techniques are equally effective, we could use whichever technique is less tedious and more enjoyable. There may also be practical issues. Although free high quality Double Entry systems are available (e.g., Barchard & Pace, 2008, 2010; Beaty, 1999; Harris, Taylor, Thielke, Payne, Gonzalez, & Conde, 2009; Lauritsen & Bruus, 2008), researchers may not be aware of them or may have difficulty using them; similarly, they may have difficulty finding and using speech synthesis software. It might turn out that Read Aloud with two people is more practical than the alternatives. Thus, research is needed on these practical issues, too.

References

Barchard, K.A. & Pace, L.A. (2008). Meeting the challenge of high quality data entry: A free double-entry system. *International Journal of Services and Standards, 4,* 359-376.

Barchard, K.A. & Pace, LA. (2010, April). *Poka-Yoke Data Entry System Version 1.19.* Excel file that allows double-entry data entry for any number of measures and items.

Beaty, J. C. (1999). The PowerChecker: A Visual Basic program for ensuring data integrity. *Behavior, Research Methods, Innovation, & Computers, 31*, 737-740.

Burchinal, M., & Neebe, E. (2006). Best practices in quantitative methods for developmentalists: 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

Harris, P. A., Taylor, R. Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics, 42*(2), 377-381.

Lauritsen, J. M. & Bruus, M. (2008). EpiData Entry (version v2.1). *A comprehensive tool for validated entry and documentation of data.* Odense, Denmark: The EpiData Association. Available at http://www.epidata.dk Accessed Nov 29, 2010.

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

Nihei, Y., Terashima, M., Suzuki, I., & Morikawa, S. (2002). Why are four eyes better than two? Effects of collaboration on the detection of errors in proofreading. *Japanese Psychological Research*, *44*(3), 173-179. doi:10.1111/1468-5884.00020

Winkler, W. E. (2004). Methods for evaluating and creating data quality. *Information Systems, 29*(7), 531-550. doi:10.1016/j.is.2003.12.003