Visual Checking – The Road to Disaster

Yevgeniya Verenikina, Cyndy Anang, Tyler Jenkin, Kelly E. Grob, and Kimberly A. Barchard
University of Nevada, Las Vegas

Reference: Verenikina, Y., Anang, C., Jenkin, T., Grob, K.E., Barchard, K.A. (2012, April). *Visual checking – the road to disaster*. Poster presented at the Western Psychological Association Annual Convention, San Francisco, CA.

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, barchard@unlv.nevada.edu

Abstract

Data entry errors can drastically affect the results of statistical analyses: They can turn strong correlations into weak ones, make significant t-tests non-significant, and reduce the values of coefficient alpha. It is therefore important to use data checking techniques that allow researchers to detect and correct errors. The first purpose of this paper was to compare the accuracy rates for three data checking techniques: double entry, read aloud, and visual checking. Each of 38 undergraduate participants pretended to be a research assistant who was checking the data for a short study using Microsoft Excel. After they finished checking the data, we compared the Excel sheet to the correct data, to determine how many errors remained. Visual checking had significantly more errors than double entry and read aloud. The second purpose of this paper was to examine the effect of data entry errors on correlations, t-tests, and coefficient alphas. Using the data that had been checked by each participant, we calculated 14 statistics (four values of coefficient alpha, four t-tests, and six correlations). We determined if each of those 14 statistics was correct, by comparing the value we obtained using the data that the participant had checked to the value that would be obtained using the data that was actually on the paper data sheets. For three of the 14 statistics, there was a significant difference: visual checking resulted in the most errors. We conclude that visual checking should not be used. If researchers use it, their published research results may be wrong. Researchers should use double entry or read aloud instead.

Introduction

Whenever researchers enter data on computers, errors are bound to happen. Unfortunately, a single data entry error can have a severe impact on statistical results (Burchinal & Neebe, 2006). Data entry errors can reduce the reliability of psychological tests. They can also reduce the strength of correlations or make a significant t-test non-significant. Even one or two data entry errors can completely alter and invalidate a statistical analysis (Velleman & Hoaglin, 1995; Wilcox, 1998).

It is important to use data checking techniques that allow researchers to identify and correct data entry errors. Several data checking techniques exist. In the visual checking technique, the data that have been entered are visually compared to the original data on the paper data sheets. In the read aloud technique, the original data are read aloud by one person while another person checks that this matches the data that have been entered. In the double entry technique, data are entered a second time and then the computer identifies any mismatches and values that are outside the allowable range. In all techniques, once an error has been identified, the researcher corrects the error.

Previous studies have found that double entry leads to fewer errors than read aloud (Kawado, Hinotsu, Matsuyama, Yamaguchi, Hashimoto, & Ohashi, 2003) and visual checking (Barchard & Pace, 2011; Reynolds-Haertle & McBride, 1992). However, no research has compared these three techniques in a single study. Thus, the first purpose of this study is to compare the accuracy rates resulting from the application of these three data checking techniques. Second, no research has examined the effect of data entry errors on statistical results. Therefore, the second purpose of this paper is to examine the effect of data entry errors on correlations, t-tests, and coefficient alphas.

Method

Participants

A total of 38 undergraduate students (20 female, 18 male) participated in the study in return for course credit. They had a mean age of 22.3 (SD = 5.26) with the ages ranging from 18 to 39. Participants reported their ethnicity as follows: Caucasian 34.2%, Asian 28.9%, Hispanic 18.4%, African-American 10.5%, Pacific Islander 5.3%, and Other 2.6%.

Procedures

Data were collected during 90-minute sessions in which participants were supervised by a trained administrator. First, participants watched a short instructional video about Microsoft Excel. Second, participants watched a randomly assigned instructional video about one of the three data checking techniques: visual checking, read aloud, or double entry. After watching both instructional videos, participants checked five data sheets (see Figure 1) to practice the assigned data checking technique. During this practice, the administrator ensured that participants used the technique correctly and answered any questions participants had. Finally, participants checked 20 additional data sheets.

Data Analysis

The first purpose of this paper was to determine if different data checking techniques result in different accuracy rates. This analysis proceeded in two steps. First, we examined the Excel files that each participant had checked, to determine how many errors they contained. An entry was considered an error if it did not match the paper data sheet. Second, we compared the accuracy rates for the three data checking techniques using a non-parametric alternative to ANOVA: the Kruskal-Wallis technique.

The second purpose of this paper was to determine if the statistical results of a study are more likely to be wrong when some data checking techniques are used. To answer this question, our analysis used three steps. First, for every participant in the sample, we used the Excel data file that they had checked in order to calculate 14 statistics: the values of coefficient alpha for each of the four scales, t-tests that compared men and women for each of the four scales, and the six correlations between all possible pairs of scales. Second, for every participant in the sample, we determined if each of those 14 statistics was correct, by comparing the value we obtained using the data that they checked to the value that would be obtained using the data that was actually on the paper data sheets. If the value was

Figure 1
Example Data Sheet

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

correct, we scored it 1. If it was incorrect, we scored it 0. Third, we compared the three data checking techniques, to determine if some techniques were more likely to result in incorrect values for correlations, t-tests, and coefficient alphas. We used the Kruskal-Wallis technique to compare the three groups for each of the 14 statistics.

Results

There was a significant difference between the distribution of errors in the double entry, read aloud, and visual checking techniques (Kruskal-Wallis chi-square(2) = 14.07, $p = .001$). In double entry, the average number of errors was 0.47; in read aloud, it was 1.14; and in visual checking, it was 12.42. See Figure 2. Pairwise comparisons showed that visual checking was significantly less accurate than double entry and read aloud. There was no significant difference between double entry and read aloud.

In our study, data entry errors led to some incorrect values for coefficient alpha, correlations, and independent sample t-tests. For three of the 14 statistics examined, there was a significant difference in the frequency of these errors across the three data checking techniques. In each of those three cases, visual checking resulted in more errors than double entry or read aloud. See Table 1.

Figure 2
Average Number of Errors for Each Data Checking Technique

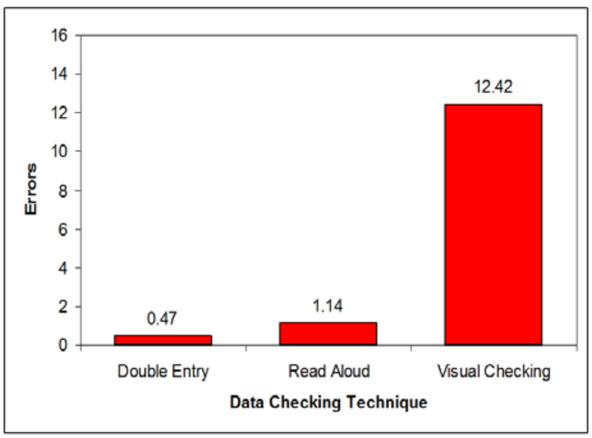

Table 1
Participants with incorrect statistics in three data checking techniques

Statistic	Number of participants with incorrect statistics: Double entry	Read aloud	Visual checking	*p*-value
Coefficient alpha				
Learning Style	3	0	0	.205
Study Habits	0	0	0	1.000
Spelling Test	1	1	1	.755
Math Test	0	2	3	.059
T-test				
Learning Style	3	0	4	.187
Study Habits	0	1	3	.087
Spelling Test	1	1	1	.755
Math Test	0	2	4	.030
Correlation				
Study habits and learning style	1	1	5	.041
Spelling test and learning style	3	1	2	.991
Spelling test and study habits	1	2	1	.228
Math test and learning style	4	2	5	.477
Math test and study habits	0	3	5	.007
Math test and spelling test	1	2	4	.115

Note. T-tests compared men and women.

Discussion

The purpose of this study was to compare three methods of data checking. Visual checking resulted in significantly more errors than double entry and visual checking, and was more likely to result in incorrect statistical results when the data were analyzed. Researchers should not use techniques that are known to invalidate their statistical findings and research conclusions. Therefore, visual checking should be abandoned.

In this study, read aloud and double entry had similar accuracy rates. Either of these two techniques could be used. Some researchers might prefer read aloud, because research assistants might prefer to work in pairs and no special software is required. Other researchers might prefer double entry, because it can be done by one person and because a record is left in the dataset to prove that errors were corrected. Free double entry systems are available as stand alone programs (Lauritsen & Bruus, 2008), Internet-based systems (Harris, Taylor, Thielke, Payne, Gonzalez, & Conde, 2009), and add-ons for Microsoft Access (Beaty, 1999) and Microsoft Excel (Barchard & Pace, 2008; 2010).

Future research should examine additional data checking techniques, to determine if there are other techniques that can be recommended, or if there are other techniques that should be avoided.

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. doi:10.1504/IJSS.2008.020053

Barchard, K. A. & Pace, L. A. (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. Available from Kimberly A. Barchard, University of Nevada, Las Vegas, barchard@unlv.nevada.edu

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.* Published online May 4 2011. doi:10.1016/j.chb.2011.04.004

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.

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,* 560-569. doi:10.1016/S0197-2456(03)00089-8

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.

Reynolds-Haertle, R. A., & McBride, R. (1992). Single versus double data entry in CAST *Controlled Clinical Trials, 13,* 487-494. doi:10.1016/0197-2456(92)90205-E

Velleman, D. C., & Hoaglin, P. F. (1995). A critical look at some analyses of major league baseball salaries. *American Statistician, 49,* 277-285.doi:10.2307/2684201

Wilcox, R. R. (1998). How many discoveries have been lost by ignoring modern statistical methods. *American Psychologist, 53,* 300-314. doi:10.1037/0003-066X.53.3.300