¹ Outrageous observations: The redheaded stepchild of data analysis

² Erin M. Buchanan[1] & Ernst-August Doelle[1,2]

³ [1] Wilhelm-Wundt-University

⁴ [2] Konstanz Business School

⁵ Author Note

⁶ Add complete departmental affiliations for each author here. Each new line herein

⁷ must be indented, like this line.

⁸ Enter author note here.

⁹ Correspondence concerning this article should be addressed to Erin M. Buchanan,

¹⁰ Postal address. E-mail: my@email.com

11   Abstract

12   Enter abstract here. Each new line herein must be indented, like this line.

13   *Keywords:* keywords

14   Word count: X

¹⁵   Outrageous observations: The redheaded stepchild of data analysis

¹⁶   *What are outliers.* Throughout psychology's assessments of experimental procedures,

¹⁷ participant observations have intrigued, inspired, and confused experimenters world wide.

¹⁸ Some of these observations exist in such extremes that we begin to think of them as outliers,

¹⁹ influential observations, or fringliers (Osborne and Overbay (2004); Wainer (1976)). Since

²⁰ they were first commented on in 1777 by Bernoulli, outliers have been defined as many

²¹ things. Muñoz-Garcia, Moreno-Rebollo, and Pascual-Acosta (1990) have synthesized their

²² own definition that we feel is concurrent with our work: "An outlier is an observation which

²³ being atypical and/or erroneous deviates decidedly from the general behavior of

²⁴ experimental data with respect to the criteria which is to be analyzed on it." (pg 217)

²⁵   Outliers can be observed in many forms, which should be outlined to ensure proper

²⁶ understanding and handling of such observations. Researchers have separated outliers into

²⁷ categories in many different ways over the years (Beckman and Cook (1983a); Hodge and

²⁸ Austin (2004); Muñoz-Garcia et al. (1990); ORR, SACKETT, and DUBOIS (1991); Osborne

²⁹ and Overbay (2004)). Many of the categories used to describe outliers have noticeable

³⁰ overlap. For instance, Muñoz-Garcia et al. (1990) stated that outliers come about by errors

³¹ in the way we gather data (inappropriate techniques, or experimenter error), errors in

³² preparation (improper hypothesis, planning, or methods), or natural variability. Shortly

³³ after, ORR et al. (1991) described outliers as consisting of people being included in an

³⁴ experiment who aren't part of the population, legitimate datapoints that are interesting

³⁵ because they do not fit the expected scheme, extreme datapoints on error components,

³⁶ human error in observation/coding, and errors in data preparation. Similarly, Hodge and

³⁷ Austin (2004) state that outliers arise, "due to mechanical faults, changes in system

³⁸ behaviour, fraudulent behaviour, human error, instrument error or simply through natural

³⁹ deviations in populations." Osborne and Overbay (2004) have also specified that outliers

⁴⁰ follow from data errors, intentional or motivated mis-reporting, sampling error,

⁴¹ standardization failure, bad assumptions about distributions, or legitimate data points.

Taken together, we have delineated the following 4 types of outliers for the purpose of this paper. 1. Data entry error, experimenter error in processing, coding, or preparing data, as well as motivated misreporting. 2. Participants who do not represent the intended population. For example, a participant who uses a DVORAK or ergonomic keyboard will not be appropriate in a study investigating the typing speeds of typical QWERTY keyboard users. 3. Valid data points which bring to light interesting phenomena the study was not aimed at capturing. For instance, the 1854 cholera epidemic in London was found by Dr. John Snow to stem from a specific water well on Broad Street. While the majority of individuals infected lived very close to the contaminated well, Dr. Snow also found an instance of cholera far away from the well. When he rode out to investigate this outlying case, he found that the woman in question had someone ride to that specific well to draw water for her because she liked the taste (Vinten-Johansen, Brody, Paneth, Rachman, & Rip, 2003). 4. Natural deviations in the population.

Effects of outliers These outliers can have serious effects on data, which can lead to imprecise data analysis (Osborne and Overbay (2004)), confusing results (Stevens (1984); Yuan and Bentler (2001)), and inappropriate conclusions (Coin (2008); Stevens (1984)). By keeping outliers in a dataset, analyses are more likely to have increased error variance (depending on sample size, ORR et al. (1991)) and biased estimates (Osborne and Overbay (2004)) as well as reduced effect size and power (ORR et al. (1991); Osborne and Overbay (2004)) which can alter the results of the analysis and lead to falsely supporting (Type I error) or denying a claim (Type II error). Additionally, incorrect estimates of effect lead to misleading meta-analyses or sample size estimates for study planning. Beyond these effects on analyses and conclusions these outliers can be inspirational to researchers and to their research models as they can encourage the diagnosis, change, and evolution of a research model (Beckman and Cook (1983b)). All together, these issues caused by outliers can lead to furthering unwarranted avenues of research, ignoring important information, and creating erroneous theories, all of which serve to weaken the sciences.

69   *Determination of outliers*

70   *Report of outliers* For these reasons, we aim to stress the importance of properly

71   identifying outliers, their potential cause(s), and ways to handle the extreme responses.

72   Further, as outlined by the APA (American Psychological Association, 2010), one

73   should acknowledge these steps not only at the time of analysis, but also at the time of

74   publication. As Kruskal, Ferguson, Tukey, and Gumbel (1960) stated, no matter how we

75   treat outlying observations, "It is very important to say something about such observations

76   in any but the most summary report. At least how many observations were excluded from

77   the formal analysis, and why, should be given." (pg. 1)

78   However, outlier report rates in studies have indicated that researchers either fail to

79   acknowledge outliers or fail to report their acknowledgment of outliers (Huffcutt and Arthur

80   (1995); ORR et al. (1991)). Huffcutt and Arthur (1995) looked at Organizational

81   Behavior/Human Resource Management meta-analytic literature since 1987. As they were

82   dealing with meta-analyses, the outliers they were identifying consisted of study coefficients,

83   or correlations, not individual data points, as outliers are usually thought of. In their study

84   they found only six of fifty meta-analyses used any outlier treatment for study coefficients.

85   Also, in most of these, highly subjective means were used to determine which study

86   coefficients were outliers. Similarly, ORR et al. (1991) inspected 100

87   Industrial/Organizational Psychology personnel studies and found no mention of outliers.

88   Although the APA and many researchers (ORR et al. (1991), . . .) have instructed

89   researchers to outline their data cleaning, there has been a dearth in recent literature

90   investigating if these instructions are actually being followed. Therefore, the authors felt that

91   the current study examined the results in 27 journals and 672 articles across 13 psychological

92   disciplines to assess if outliers are being reported. Given previous report rates (Huffcutt and

93   Arthur (1995); ORR et al. (1991)) we expect to find little to no reporting of outliers in these

94   articles. Further, we will examine how outlier reports relate to field, journal, analysis, and

95   sample size. We will also investigate the methods and rational of outlier detection and

96 removal.

Method

Fields

99 A list of psychological field areas was created to begin the search for appropriate

100 journals to include. The authors brainstormed the list of topics (shown in Table XX) by first

101 listing major research areas in psychology (i.e. cognitive, clinical, social). Second, a list of

102 common courses offered at large universities was consulted to add to the list of fields. Lastly,

103 the American Psychological Association's list of divisions was examined for any potential

104 missed fields. The topic list was created to capture large fields of psychology with small

105 overlaps (i.e. cognition and neuropsychology) while avoiding specific subfields of topics

106 (i.e. cognition, perception, and memory).

Journals

108 Once these topic areas were decided, researchers used various search sources (Google,

109 EBSCO host databases) to find journals that were dedicated to each broad topic. Journals

110 were included if they appeared to publish a wide range of articles within the selected fields. A

111 list of journals, publishers, and impact factors (as noted by the journal website) was created

112 for each field. Two journals from each field were selected based on the following criteria: 1)

113 high impact factors, 2) impact factors over one a minimum, 3) a mix of publishers if possible,

114 and 4) availability due to university resources. These journals are shown in Table XX.

Articles

116 Fifty articles from each journal were examined for data analysis. Data collection of

117 articles started at the last volume publication from 2012 and progressed backwards until fifty

118 articles had been found. We excluded online first publications and started in 2012 to ensure

119 time for errata and retraction of articles. Articles were including if they met the following

criteria: 1) included data analyses, 2) included multiple participants or data-points, and 3)
analyses were based on human subjects or stimuli. Therefore, we excluded theory articles,
animal populations, and single subject designs. Based on article review, three fields were
excluded. Applied behavior analysis articles predominantly included single-subject designs,
evolutionary psychology articles were primarily theory articles, and statistics related journal
articles were based on user created data with specific set characteristics. Since none of these
themes fit into our analysis of understanding data screening with human subject samples, we
excluded those two fields from analyses.

Data Processing

Each article was then reviewed for key components of data analyses. Each experiment
in an article was treated separately. For each experiment, the type of analysis, number of
participants/stimuli, and if they indicated outliers were coded. Types of analyses were
broadly coded into basic statistics (descriptive statistics, z-scores, t-tests, and correlations),
ANOVAs, regressions, chi-squares, nonparametric statistics, modeling, and Bayesian/other
analyses. First, we coded if outliers were mentioned at all in article. If so, we coded outliers
into four types: 1) people, 2) data, 3) both, and 4) none found. We found that a separate
coding for data was important for analyses with response time studies where individual
participants were not omitted but rather specific data trials were eliminated. Then, the
author decision on what to do with the outliers was coded into whether they removed
participants/stimuli or left these outliers in the analysis, as well as if they tested the analyses
with and without these outliers for determination of their effect on the study. If they
removed outliers, a new sample size was recorded. Lastly, we coded the reasoning for outlier
detection as either participant based (i.e. 3-month old infants were too fussy to be included),
experimenter based (i.e. the experimental session was interrupted), statistical based
(i.e. three SD from the mean), and no listed reason.

Table XX ##List of Fields, Journals, and Impact Factors 2012

Field	Journal
Applied Behavior Analysis	Journal of Experimental Analysis of Behavior
Applied Behavior Analysis	Journal of Applied Behavior Analysis
Clinical	Journal of Consulting and Clinical Psychology
Clinical	Journal of Clinical Psychology
Cognitive	Cognitive Psychology
Cognitive	Journal of Experimental Psychology: Learning, Memory, and Cognitio
Counseling	Journal of Counseling
Counseling	The Counseling Psychologist
Developmental	Journal of Experimental Child Psychology
Developmental	Journal of Youth and Adolescence
Educational	Journal of Educational Psychology
Educational	Contemporary Educational Psychology
Environmental	Journal of Environmental Psychology
Environmental	Environment and Behavior
Evolutionary	Evolution and Human Behavior
Evolutionary	Evolutionary Psychology
Forensics	Psychology, Public Policy, and Law
Forensics	Law and Human Bevhavior
Industrial Organization	Organizational Behavior and Human Decision Process
Industrial Organization	Personnel Psychology
Neurological/Physiological	Neuropsychology
Neurological/Physiological	Cognitive, Affective, and Behavioral Neuroscience
Social	Journal of Personality and Social Psychology
Social	Journal of Experimental Social Psychology
Sports	Journal of Sport & Exercise Psychology
Sports	Sociology of Sport Journal
Statistics	Special Section of the Psychological Bulletin
Statistics	Structural Equation Modeling

146

147   *Note.* Impact factors as of tme of data collection (Spring 2013).

148 **Participants**

149 **Material**

150 **Procedure**

151 **Data analysis**

152 **Results**

153 **Discussion**

References

Beckman, R. J., & Cook, R. D. (1983a). Outlier.........s. *Technometrics*, *25*(2), 119. doi:10.2307/1268541

Beckman, R. J., & Cook, R. D. (1983b). [Outlier.........s]: Response. *Technometrics*, *25*(2), 161. doi:10.2307/1268548

Coin, D. (2008). Testing normality in the presence of outliers. *Statistical Methods and Applications*, *17*(1), 3–12. doi:10.1007/s10260-007-0046-8

Hodge, V. J., & Austin, J. (2004). A survey of outlier detection methodologies. *Artificial Intelligence Review*, *22*(2), 85–126. doi:10.1007/s10462-004-4304-y

Huffcutt, A. I., & Arthur, W. (1995). Development of a new outlier statistic for meta-analytic data. *Journal of Applied Psychology*, *80*(2), 327–334. doi:10.1037//0021-9010.80.2.327

Kruskal, W., Ferguson, T. S., Tukey, J. W., & Gumbel, E. J. (1960). Discussion of the Papers of Messrs. Anscombe and Daniel. *Technometrics*, *2*(2), 157–166. doi:10.1080/00401706.1960.10489890

Muñoz-Garcia, J., Moreno-Rebollo, J. L., & Pascual-Acosta, A. (1990). Outliers: A formal approach. *International Statistical Review / Revue Internationale de Statistique*, *58*(3), 215–226. doi:10.2307/1403805

ORR, J. M., SACKETT, P. R., & DUBOIS, C. L. Z. (1991). Outlier detection and treatment in I / O psychology: A survey of researcher beliefs and empirical illustration. *Personnel Psychology*, *44*(3), 473–486. doi:10.1111/j.1744-6570.1991.tb02401.x

Osborne, J. W., & Overbay, A. (2004). The power of outliers (and why researchers should always check for them). *Practical Assessment, Research & Evaluation*, *9*(6), 1–12.

Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychological Bulletin*, *95*(2), 334–344. doi:10.1037/0033-2909.95.2.334

Wainer, H. (1976). Robust statistics: A survey and some prescriptions. *Journal of*

180   *Educational Statistics*, *1*(4), 285. doi:10.2307/1164985

181   Yuan, K. H., & Bentler, P. M. (2001). Effect of outliers on estimators and tests in covariance

182   structure analysis. *British Journal of Mathematical and Statistical Psychology*, *54*(1),

183   161–175. doi:10.1348/000711001159366