Hammill Institute on Disabilities

Remedial and Special Education
2018, Vol. 39(3) 171–176
© Hammill Institute on Disabilities 2018
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0741932517745491
rase.sagepub.com
SAGE

Establishing Journalistic Standards for the Publication of Negative Results

Angus Kittelman, MS[1], Cody Gion, MS[1], Robert H. Horner, PhD[1], Joel R. Levin, PhD[2], and Thomas R. Kratochwill, PhD[3]

Abstract

We argue in this article that there are conditions in which publication of negative results can make a useful contribution. Three small-scale examinations of journal publication criteria for publishing negative results were conducted. We first reviewed 29 journals from education and school/counseling psychology to assess author submission guidelines for reporting negative results. We then reviewed the most recent issue of each of the 29 journals to determine whether articles reporting negative results were being published. Finally, we surveyed the editors of the 29 journals to examine conditions in which they perceived it appropriate to publish articles that did not demonstrate experimental effects. Results of these studies indicate that currently only one of the 29 journals provides formal guidance to authors about submitting papers with negative results. Two articles published in the 29 journals in their last issue of 2016 contained only negative results. And, of the 60% of the editors who responded to our survey, 96% indicated there were conditions in which publication of negative results was appropriate. In conclusion, we propose situations in which studies reporting negative results would make a valuable contribution.

Keywords

journalistic standards, negative results

The purpose of this article is to encourage journalistic standards for publication of negative results in professional journals. Negative results (also known as "null results") are associated with experimental studies in which an administered treatment or intervention does not produce an experimentally convincing demonstration of the intended or desired outcome(s). Seftor (2016) recently outlined three reasons why intervention studies may fail to demonstrate meaningful effects. First, negative results may reflect a failure of active ingredients or theoretical principles (i.e., the intervention was ineffective). Second, negative results may be due to poor implementation (e.g., low treatment integrity). Third, negative results may occur even if the intervention was effective but the effect was not adequately captured properly because of psychometric or statistical inadequacies (e.g., unreliable measure or low statistical power).

The premise for this article is that there are conditions under which negative results provide a useful contribution to the field (e.g., Ferguson & Heene, 2012; Knight, 2003). If this premise is accepted, then review of the publication standards and practices of major professional journals may serve to document how studies describing negative results are being utilized, which in turn will provide guidance for the future. We argue that studies reporting negative results can be valuable additions to a research domain's cumulative knowledge basis. Our argument is based on (a) a review of the published submission guidelines of 29 highly ranked educational and psychological journals, (b) an examination of the rate of published articles reporting negative results in the last issue from 2016 for each journal, and (c) a survey of the opinions of journal editors from these journals about whether and under what conditions publication of articles reporting negative results is appropriate. We conclude by recommending how journal editorial boards may productively encourage, review, and consider the publication of intervention-research studies that contain negative results.

The Contribution of Treatment/Intervention Studies Reporting Negative Results

Science is a formal process for identifying the structure and behavior of the physical and natural world. Scientific knowledge helps us describe, predict, and control phenomena. The

[1]University of Oregon, Eugene, USA
[2]University of Arizona, Tucson, USA
[3]University of Wisconsin–Madison, USA

Corresponding Author:
Robert H. Horner, University of Oregon, 1235 University of Oregon, Eugene, OR 97403, USA.
Email: robh@uoregon.edu

use of experimental procedures to identify scientific knowledge has been the source of unending discussion and debate and yet, no enduring set of core standards has emerged (T. D. Cook & Campbell, 1979; Shadish, Cook, & Campbell, 2002). The standards currently used to select peer-reviewed research reports provide one operational index of how science is defined within any discipline. Within many areas of education and psychology, it is common for reviewers and editors to expect an intervention study to (a) incorporate a level of experimental control through application of traditional group or single-case research designs that allow causal inference related to any observed changes in desired outcomes, (b) provide replicable description of research procedures, (c) examine one or more outcome measures with acceptable reliability and validity, and (d) document change in at least one valued outcome (Gersten et al., 2005; Kratochwill et al., 2010; Odom et al., 2005).

Even if the reviewer expectations listed above are impeccably satisfied, when intervention-research results document no change in valued outcomes, publication becomes more difficult. The question for reviewers and editors lies in identifying how the results from a study make a useful contribution to the field. Related to the just-presented reviewer expectation, Kratochwill, Levin, and Horner (2017) propose that the publication of intervention studies producing negative results may make useful contributions if four criteria are met.

- First, a study should measure a valued phenomenon with adequate reliability and construct validity. The focus of the study should be on a valued outcome, and unless the data being reported can be believed to represent the targeted phenomenon, any scientific inference is unwarranted.
- Second, a study should document that any manipulation of an independent variable (i.e., delivery of an intervention) occurred as intended, thereby exhibiting intervention integrity (or fidelity). Research procedures should be described with replicable precision and with documentation that allows readers to be confident about what was done (Sanetti & Kratochwill, 2014).
- Third, a study should incorporate a research design that controls for plausible threats to the study's internal validity (Kratochwill et al., 2010; Shadish et al., 2002).
- Finally, the study's results need to be placed in professional context. For example, documenting that an intervention or procedure did not produce an intended or desired effect (i.e., convincingly demonstrating that a negative result occurred) becomes an important contribution if the study is viewed in conjunction with related studies that have documented conditions under which positive results were demonstrated (e.g., with somewhat different intervention variations or dosage levels, with different participant samples, in different situational contexts). Insofar as few interventions or procedures work under all conditions, documenting where valued outcomes are likely to occur and where they are unlikely to occur is of both conceptual and clinical value.

We propose that although not all studies reporting negative results make a valuable contribution to their respective fields, a substantial proportion of studies that meet Kratochwill et al.'s (2017) four criteria warrant formal publication. Working from this assumption, we conducted three small-scale examinations of journal publication criteria and standards within the fields of education and psychology. The first examination involved analysis of the author submission guidelines provided by 29 journals, the second was a review of articles reporting negative results in those same 29 journals using each of their last issues of 2016 (which ranged from September 2016 to December 2016), and the third was an online survey of journal editors to determine their predilections about whether and when publication of studies with negative results would be appropriate.

Author Submission Guidelines

Method for Study 1

To determine whether journalistic standards already exist for publishing studies reporting negative results, we reviewed journal submission guidelines from the top 10 ranked peer-reviewed journals in the fields of general education, special education, educational psychology, and counseling. Journal rankings were based on Google Scholar metrics and ordered by their 5-year h-index and h-median metrics, which cover articles published between 2011 and 2015 and which were last updated in June 2016. Because one journal overlapped in two separate fields (i.e., general education and educational psychology/counseling), we reviewed a total of 29 journal submission guidelines (a list of the 29 journals reviewed may be obtained from the authors on request). Each journal submission guideline was coded as "yes" if the journal guidelines included standards for publishing studies that report negative results or "no" if the journal guidelines did not include standards for publishing studies that report negative results." For interrater agreement (IRA), eight of the 29 submission guidelines (27.6%) were reviewed by two of the present authors independently. IRA was calculated by dividing the total number of journal guidelines on which there could be agreement by the number of agreements, and was found to be 8 out of 8 (100%).

Results for Study 1

From the 29 journal submission guidelines reviewed, only one contained a passage that discussed (and in fact "welcomed")

the submission of studies that may report negative results. Specifically, the passage states,

> Since researchers, policymakers and practitioners need to know how the possibilities of educational technology can be maximised and problems of adoption and sustainability minimised, BJET [British Journal of Educational Technology] particularly welcomes such submissions as . . . reports on educational technology initiatives which met with problems and/or failed to achieve their aims, and the lessons to be learned from these.

Published Articles Describing Negative Results

Method for Study 2

To examine the rate at which intervention studies reporting negative results were published in these 29 journals, we first reviewed abstracts in the journals' last issues of 2016 (between September 2016 and December 2016). All abstracts were screened to determine whether they met the initial inclusion criteria of being an intervention study (i.e., experimental group, quasi-experimental group, and single-case). We then screened all articles that met the initial abstract screening criteria of being an intervention study to determine whether articles contained negative results, partial (i.e., mixed), or only positive findings. We screened a total of 285 abstracts from the 29 journals (the number of abstracts ranged from 4 to 42 per the journals' issues). Of the 285, 55 (19%) were identified as intervention studies and fully reviewed. Two abstracts for each of the 29 journals (n = 58, 20% of all abstracts) were randomly selected using a random number generator and coded for IRA by the second author. IRA was 95% (55 out of 58) for the abstracts screened. We then reviewed the three abstract coding discrepancies until we reached 100% agreement. Next, we conducted full article IRA screening to determine whether articles contained mixed/negative results effects or only positive effects by randomly selecting 15 of the 55 articles (27%) identified as intervention studies. IRA was 87% (13 out of 15) on the articles that were fully screened. Finally, we reviewed the full screening discrepancies until 100% agreement was reached.

Results for Study 2

Of the 55 intervention studies, 23 of the 55 (42%) were identified as having mixed findings. Only 2 of the 55 (4%) were found to contain only negative results (as defined by "statistically nonsignificant" or "no functional relation"), with no positive (i.e., "statistically significant" or "functional relation") findings reported. Of the 23 studies identified as having mixed findings, 18 (78%) were experimental group or quasi-experimental group studies, four (17%) were identified as using mixed methods (i.e., group quantitative and single-case methods), and two of the 23 studies (9%) were identified as single-case intervention studies. Of the 2 studies found to contain only negative results (9%), both were identified as experimental group or quasi-experimental group studies.

Survey of Journal Editors' Perceptions

Method for Study 3

To examine whether studies reporting negative results are perceived as appropriate for publication, we developed a brief online survey and recruited the executive editors of the 29 aforementioned journals using contact information provided on each journal's website. The survey questions were designed to explore under what conditions the editors would be willing to consider publishing studies that reported negative results. The survey was created by the authors of this article, and was sent to reviewers with extensive research and editorial experience within the fields of psychology and general and special education for their opinions and comments. The survey included an introductory letter and four questions each, with the response options "no," "maybe," and "yes," followed by space for comments. The three "primary" survey questions assessed whether the editor would (1) "support publication of a study that employs rigorous measurement and design procedures with high treatment integrity but does not demonstrate an experimental effect (e.g., statistical results are nonsignificant, small effect size, or single-case results fail to demonstrate a functional relation)"; (2) "support publication of a study that fails to demonstrate an experimental effect for a previously published (or commonly used) practice, even though the study employs rigorous measurement and design procedures with high treatment integrity"; and (3) "support publication of a study that fails to demonstrate an experimental effect but also reports poor treatment integrity with professionally rigorous measurement and design procedures."

In addition, a fourth question was included in the survey to explore whether editors would support an editorial policy originally proposed by Rosenthal (1966, p. 36) and echoed by Walster and Cleary (1970) to help counteract the well-documented concerns about "publication bias" and the associated "file drawer problem" (Rosenthal, 1979). For related discussion, see Kratochwill et al. (2017), Tincani and Travers (2017), and B. G. Cook and Therrien (2017). In our survey's context, publication bias focused on the issue of reviewers and editors of experimental-research journals being more positively disposed toward submitted manuscripts that report positive (statistically significant) results than submitted manuscripts that report negative results (statistically nonsignificant) results, which of course leads to a

much greater editorial rejection—and ultimately nonpublication—of the latter than of the former manuscript types.

Rosenthal's (1966) suggested change in current editorial policy (which persists today, half a century later) would be for authors initially to submit manuscripts for review that *include* the study's introduction, rationale, and method focusing on the study's participants, design, materials, procedures, and outcome measures, but that *omit* anything pertaining to the study's results and discussion of those results. Such a procedure would essentially result in a review of the quality of a "data-less" study's conceptualization and methodology, as well as its potential value, with the benefit that reviewers' and editors' publication recommendations would not be unduly influenced by the obtainment of positive results. Because this question went beyond the primary intent of the survey, it was made clear to the surveyed editors that responding to the question was optional.

The question was presented to editors as follows: "Over the years, there have been suggestions designed to counteract the well-documented 'publication bias' on the part of journal reviewers and editors (i.e., their favoring acceptance of studies that report statistically significant results). One recurring suggestion is that when authors submit their research-based manuscripts for review, the reports include the introduction, the literature review, the study's rationale and hypotheses, and the method sections but that they omit the results and discussion sections. Assuming that the initial 'data-less' manuscript is deemed 'worthy' by the reviewers and/or editor, the results and discussion sections would then be submitted for editorial review—with the understanding that a 'worthy' study would eventually be recommended for publication as long as the data were appropriately analyzed/reported and the discussion acceptably presented. Such a suggestion embodies Robert Rosenthal's (1966) concern about the nonpublication of negative results, an already recognized issue of publication bias at the time, and which subsequently led to what he termed the 'file drawer problem' (Rosenthal, 1979):

> What we need is a system for evaluating research based only on the procedures employed. If the procedures are judged appropriate, sensible, and sufficiently rigorous to permit conclusions from the results, the research cannot then be judged inconclusive on the basis of the results and rejected by the referees or editors. Whether the procedures were adequate would be judged independently of the outcome. To accomplish this might require that procedures only be submitted initially for editorial review or that only the result-less section be sent to a referee or, at least, that an evaluation of the procedures be set down before the referee or editor reads the results. (Rosenthal, 1966, p. 36)

Would you support such a policy in some form, or with certain modifications?"

Because 11 of the 29 journals had multiple executive editors, a total of 47 editors were initially invited to participate in the survey. In addition, five co-executive editors from one of the 29 journals notified the authors that they had completed the survey as a group, which reduced the number of editors who independently completed the survey to 43. Response rate was calculated by dividing the total number of respondents who completed the survey by the total number of respondents who were invited to participate in the survey. A total of 26 out of 43 (60%) editors completed at least one of the survey questions.

Results for Study 3

Of the 43 responding editors, 33 (77%) editors opened the survey, 26 (60%) completed Question 1, 24 (56%) completed Question 2, 23 (53%) completed Question 3, and 21 completed the optional Question 4.

For Question 1, 25 of the 26 editors (96%) indicated they would consider publishing studies that used rigorous measurement and design procedures with high intervention integrity but without experimental effects. Sixteen of the 25 (64%) indicated they would stipulate conditions for publication (i.e., those responding "maybe"). Only one of the 26 editors (4%) responding to Question 1 indicated that he or she would not consider publishing manuscripts with these characteristics.

For the 16 editors indicating "conditions" for publication, the stipulated conditions included are paraphrased as follows: (a) whether the study would make an important contribution to advancing the literature/field (n = 3); (b) addressing an unanswered question in a particular field (n = 2); (c) reporting a study that runs counter to findings of a long-standing and/or established area of work that calls into question a new area of research (n = 1); (d) providing a strong literature review, theoretical framework, and discussion acknowledging the lack of intervention effects with implications (n = 1); (e) utilizing procedures adopted by other researchers (n = 1); (f) the researcher applying another intervention to show that the dependent variable is amenable to change (n = 1); (g) if a careful examination of the study indicated top-notch methodology and quantitative analyses (n = 1); and (h) whether the study had reviewer support (n = 1).

All 24 of the editors who completed Question 2 indicated they would be willing to publish studies reporting negative experimental results if the intervention under review had been documented to be effective in earlier reports and the study employed rigorous measurement and design procedures with high treatment integrity (i.e. high-quality nonreplication study). Thirteen of the 24 editors (54%) stipulated conditions that would encourage publication, with these conditions including (a) the negative results overturned a long-held finding/belief (in which previous findings and/or established interventions were supportive of these findings; n = 3), (b) the quality of the author's case or

argument was convincing enough to explain/suggest why the nonreplication occurred ($n = 2$), (c) if the attempted replication study incorporated a comparable sample and procedures to the original study ($n = 2$), (d) if the study made an important contribution to the literature, added value to the field, or influenced practices ($n = 2$), (e) if the author(s) identified a way to impact a particular behavior and proved that it was amenable to change ($n = 1$), (f) if the study adopted quantitative methodology ($n = 1$), and (g) if the study had reviewer support ($n = 1$).

Question 3 described conditions under which a study did not produce positive results but adopted rigorous measurement and design procedures with low intervention integrity. Only nine of the 23 editors (39%) responding to this question said they would consider publishing such an article, or would publish it under certain conditions. The remaining 14 editors (61%) responding to Question 3 indicated that they would not publish studies with negative results based on questionable intervention integrity. Open-ended paraphrased responses from the five editors indicating they "may under certain conditions" publish such studies focused on whether (a) the study was quantitative in nature ($n = 1$), (b) the study had reviewer support ($n = 1$), or (c) the argument/justification that was made by the researchers was convincing as to why the study design was not implemented as planned ($n = 1$).

Question 4 examined the openness of editors to a modified ("data-less") journal review process in which authors initially submit only their study introduction, rationale, and methods. The study's results would later be submitted and examined only for those studies that received high marks for the first-stage evaluation criteria. Seventeen of the 21 editors (81%) responding to this question indicated they would or might (under certain conditions) support this review process, while four of the 21 editors (19%) indicated they would not support this approach. Of the 17 editors who were willing to consider this process, 14 (82%) offered comments, including the following verbatim comments from seven editors:

> It is hard to envision what exactly the papers would be evaluated on—also with a high number of submissions it would be hard to argue to publish a large number of negative results [studies] when you have many excellent papers with [statistically] significant results, [however] I have had this topic come up several times as editor, [and] right now we do not really publish negative results but I find this an important and interesting question.

> In theory, sound[s] like a good idea, [however] I think early implementation could be complex in several ways but it is an exciting idea.

> Sometimes it is difficult to ascertain the experimental rigor [of a study] without examining the data [and] I would want to make sure that there is not the opposite problem—a bias toward publication—if reviewers believe a study was deemed worthy of publication based on a previous process of review.

> Seems as if it would add [a] further step to an already rigorous process.

> There are a lot of issues here [and] often articles are rejected due to methods and/or inadequate discussion [and] in addition, there is the issue of time in review and in acceptance to publication.

> I would generally support [the idea], but not necessarily for the journal I currently edit.
> [D]epends on the extent to which the negative results impact accepted practice, current wisdom, or paradigm.

In addition to tabulating the editors' responses question by question, we conducted all possible pairwise-question comparisons of their "positive-inclination" responses (as defined by the proportion of "yes" and "maybe" responses) through statistical tests of correlated proportions. Of the six pairwise tests conducted, three of them yielded statistically significant differences in the proportion of positive-inclination responses, with not surprisingly all involving Question 3 (publishing low treatment-integrity studies with negative results): (a) A greater percentage of positive-inclination responses was found for Question 1 (publishing high treatment-integrity original studies with negative results) than for Question 3, 96% versus 39%, respectively, $p < .01$; (b) similarly, a higher percentage of editors were positively inclined to support Question 2 (publishing high integrity studies that did not replicate a previous positive finding or practice) than Question 3, 100% versus 39%, respectively, $p < .01$; and (c) a greater percentage of editors were positively disposed toward Question 4 (implementing a data-less manuscript review process) than Question 3, 81% versus 38%, respectively, $p < .05$.

Summary and Discussion

In reviewing the submission guidelines of 29 top ranked peer-reviewed journals, and the rate at which articles published in these journals report negative results, we can draw several conclusions. First, submission of studies that report negative results is not widely encouraged. Only one journal included guidelines indicating that research reporting negative results was actively encouraged. Second, our review of the journals' last issues of 2016 indicates that the publication of studies with only negative results does occur, but it is very uncommon (4%), and it is most likely to occur within a study that documents both positive experimental effects and negative results (42%).

Interestingly, when editors from these 29 journals were asked whether they would support the publication of studies, or replications of previously conducted studies, that report negative

results under certain conditions, the vast majority expressed a positive inclination to do so. The primary stipulation was that studies with negative results that warranted possible publication would be expected to adopt rigorous measurement and design procedures with high treatment integrity.

Overall, the results suggest that although the publication of negative results is not common or openly encouraged, most editors are willing to publish studies that report negative results, under specific conditions. Authors of manuscripts that report negative results will want to (a) ensure that their research is methodologically and psychometrically sound with high treatment integrity and (b) document how their results provide a valuable contribution to a particular field of study.

Although we anticipate that the publication of negative results will remain a limited voice in the field, it would be helpful to articulate the conditions and standards for which these studies are found to make a valuable contribution. With this goal in mind, we recommend adding the following passage to submission guidelines for peer-reviewed journals:

> [name of journal] welcomes the submission of manuscripts reporting intervention studies that do not produce experimentally convincing (conventionally, statistically significant) effects, as long as the reported research: (a) adopts rigorous measurement and design procedures with high intervention integrity; and (b) provides a meaningful contribution to the field.

In addition, we suggest that editors provide orientation materials to all members of a journal's editorial board indicating the conditions under which negative results may warrant publication. Conditions appropriate for publication of negative results may vary depending upon each journal's area of focus. One procedural option that editors might give serious consideration is the Rosenthal (1966) and Walster and Cleary (1970) proposition, to conduct an initial review only of the introduction, rationale, and methods of an experimental study, and reserve evaluation of results only for those manuscripts that meet initial methodological and substantive criteria.

As outlined by Kratochwill et al. (2017), under certain conditions, the publication of experimental results documenting negative results has the potential to make important—and more representative—contributions to one's specific field of inquiry (Cooper, DeNeve, & Charlton, 1997; Rosenthal, 1979). The present authors believe that having these standards explicitly stated in a journal's submission guidelines and publishing-related protocols holds promise for furthering the contributions from well-conducted scholarship.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Development of this paper was supported by the Office of Special Education Programs, U.S. Department of Education (H326S130004 and R324U140001). Opinions expressed herein are those of the authors and do not necessarily reflect the position of the U.S. Department of Education, and such endorsements should not be inferred.

References

Cook, B. G., & Therrien, W. J. (2017). Null effects and publication bias in special education research. *Behavioral Disorders*, *42*, 149–158. doi:10.1177/0198742917709473

Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis for field settings*. Chicago IL: Rand McNally.

Cooper, H., DeNeve, K., & Charlton, K. (1997). Finding the missing science: The fate of studies submitted for review by a human subjects committee. *Psychological Methods*, *2*, 447–452.

Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Psychological Science*, *7*, 555–561.

Gersten, R., Fuchs, L., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children*, *71*, 149–164.

Knight, J. (2003). Negative results: Null and void. *Nature*, *422*(6932), 554–555.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/ReferenceResources/wwc_scd.pdf

Kratochwill, T.R., Levin, J.R., & Homer, R.H. (2017). Negative results: Conceptual and methodological dimensions in single-case intervention research. *Remedial and Special Education*. Advanced online publication. doi: 10.1177/0741932517741721

Odom, S., Brantlinger, W., Gersten, R., Horner, R., Thompson, B., & Harris, K. (2005). Research in special education: Scientific methods and evidence-based practices. *Exceptional Children*, *71*, 137–148.

Rosenthal, R. (1966). *Experimenter effects in behavioral research*. New York, NY: Appleton Century Crofts.

Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*, 638–641. doi:10.1037/0033-2909.86.3.638

Sanetti, L. M., & Kratochwill, T. R. (2014). *Treatment integrity: A foundation for evidence-based practice in applied psychology*. Washington, DC: American Psychological Association.

Seftor, N. (2016). *What does it mean when a study finds no effects?* (REL 2017-265). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/edlabs

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, MA: Houghton Mifflin.

Tincani, M. & Travers, J. 2017. Publishing Single-Case Research Design Studies That Do Not Demonstrate Experimental Control. *Remedial and Special Education. Advanced online publication*. doi: 10.1177/0741932517697447

Walster, W., & Cleary, T. A. (1970). A proposal for a new editorial policy in the social sciences. *American Statistician*, *24*(2), 16–19.