

A PEER-REVIEWED PUBLICATION

The Journal of Human Sexuality

Alliance for Therapeutic Choice and Scientific Integrity

FEATURE ARTICLE

An Analysis of the Retracted Article
by Rabak and Lan (2023):
A Challenge for Peer Review
and Scientific Integrity

Journal of Human Sexuality · Vol. 15

An Analysis of the Retracted Article

by Rabak and Lan (2023):

A Challenge for Peer Review and Scientific Integrity

Walter R. Schumm, Ph.D.¹

¹Kansas State University

An article, allegedly by Rabak and Lan (2023), was retracted after its authors confessed to its research as fabricated (Rabak & Lan, 2024). Possibly the authors meant to create a “win, win” catch-22 trap situation where if the fabricated paper was accepted, it would seem to discredit a journal in which Sullins (2015) had published a controversial article. But if it was rejected, the authors could have claimed the journal was biased against papers with LGBT topics, which might also seem to discredit the journal and Sullins (2015). Notably, the editor and reviewers for Sullins (2015) were *not* the same as those for Rabak and Lan (2023), limiting whatever point was at issue. Furthermore, numerous top-tier journals have published articles later retracted, negating the idea that a journal is “fake” just because one of its articles is later retracted or had included fabricated data or results. Rabak and Lan (2023) survived peer review in part because a second round of peer review was avoided, possibly because the authors—and thus the editor—had pleaded for an extremely rapid review (three days). However, superficial errors may only be typographical and can be corrected and may not mean an article has been faked. Most editors and reviewers assume that submitted papers are legitimate because of the serious nature of fabrication and subsequent adverse consequences for scholars if detected. At the same time, there are statistical tests available for deeper testing of research that can detect fabrication, tests that did indicate that Rabak and Lan (2023) was a fabrication. Since Rabak and Lan (2024) have indicated that they had wanted to discredit Sullins (2015), a further discussion of the merits and limitations of Sullins (2015) is included. Implications for journal editors and peer reviewers are discussed, as well as for graduate education in research methodology and ethics.

Keywords: Questionable research practices, fabricated articles, retractions, scientific fraud, peer review

Walter R. Schumm, Ph.D., is Emeritus Professor of Applied Family Science, Department of Applied Human Sciences, College of Health and Human Sciences, Kansas State University, recently retired after 42 years of teaching and research at Kansas State. His earned degrees include a bachelor’s degree in physics from the College of William and Mary (Virginia), 1972; a master’s degree in family and child development from Kansas State University, 1976; and a Ph.D. in family studies from Purdue University in 1979. He was the editor of the journal *Marriage & Family Review* from 2010 to 2020 and continues to serve as an associate

editor/reviewer for numerous social science journals. He is a Fellow with the National Council on Family Relations (NCFR) and a certified family life educator. His primary research interests have included military families and same-sex parent families, often with a focus on strengths or weaknesses of methodologies used in studying diverse types of families. He retired as a colonel, U.S. Army, in 2002 after nearly 30 years of service in the Army Reserve and Army National Guard, having commanded units from detachment, separate platoon, and company levels to battalion and brigade levels.

Correspondence concerning this article should be addressed to Dr. Walter R. Schumm, 1151 Lane 6, Powell, WY 82435. Email: schumm@ksu.edu.

Because I was one of the three scholars who reviewed Rabak and Lan (2023), I would like to respond to their admission (Rabak & Lan, 2024) that their article used fabricated results (i.e., was faked). Because the authors appear to remain anonymous, I shall refer to their discussion of their fake paper and of Sullins (2015) by the pen names of Rabak and Lan (2024). Considerations of the reproducibility of science (Hergovich et al., 2010; Johnson et al., 2017; Nosek et al., 2022; Open Science Collaboration, 2015; Patil et al., 2016) and/or confirmation bias (Nickerson, 1998; Schumm, 2015, 2021) may be related to this issue but are beyond the scope of the discussion here.

History

Rabak and Lan (2023) sent their paper to the *Journal of Education, Society, and Behavioural Science* to try, it seems, to demonstrate that its review process was less than thorough, or alternatively, that it featured an anti-LGBTQ+ bias. They requested a rapid review, an option offered by the journal. The editor asked three reviewers, including myself and reviewers from India and Hong Kong, to return their reviews within three days, which we did, with suggested revisions. Specifically, the email letter dated March 6, 2023, requesting my review, said specifically: “Author of this manuscript politely requested an urgent decision due to her upcoming job interview and thesis

submission. Author of this manuscript will be grateful to the reviewer for the urgent review comments. I would be grateful if you would kindly find some time to review the above-mentioned manuscript and send your valuable comments within 8 March 2023.” This comment implied that a female graduate student needed help desperately and urgently, even though it is not clear how the editor knew the primary author was a woman or even a graduate student per se. This request was a very unusual imposition upon the prospective peer reviewers and may have set the stage for them to focus on unusually fast and possibly less rigorous reviews.

However, it appears that the editor accepted the paper before the reviewers had a chance to review a revised version of the paper (the editor, who was from Romania, did see a revised version). In my review, I noted several limitations of the paper: the paucity of references, it did not study fathers, the sampling process was unclear, scale reliabilities were not reported, main effects (handedness, religiosity, education, age, race, parental sexual orientation) in the regression analysis were not reported, social desirability response bias had not been measured or controlled, and the minimum and maximum ages of the mothers were not reported, making a test of the validity of the data impossible (since legitimate data will usually have a range of results from three to eight if you divide the range by the standard deviation; Schumm et al., 2018).

Rabak and Lan (2023) may have cringed a bit as my comment on the ages of the mothers came close to questioning if their paper might have been fabricated; they did not appear to respond to that particular comment in their reply to the editor and did not include the minimum and maximum ages of the mothers in their revision. Had I seen a revised version of the paper that left out that information on mothers' ages, it would have been a red flag, increasing my suspicions about the report. It probably should have been a red flag for the editor regardless. The other reviewers noted that the paper was only "fair" in terms of scientific correctness, that the abstract was not comprehensive, and that the paper had structural issues. It would appear that the reviewers, on first review, were looking for major issues rather than specific problems with internal consistency and might have focused on smaller issues if a revision had dealt with the "major" issues already raised. Furthermore, they naively trusted the authors to be legitimate, honest scholars, and were not assessing the paper for evidence of fabrication, even though one reviewer did hint at that possibility.

After their paper (Rabak & Lan, 2023) was accepted and published, they claimed that the publication process was not only deceitful (changed the date of receiving the paper to make the review period seem longer than it was) but "comically inept" (Rabak & Lan, 2024). Rabak and Lan (2024) noted numerous issues with their 2023 article that they believed should have come to the attention of the review process. The names and the assigned university were fake. They provided an incorrect square kilometer area for Singapore (720/500). One of their several demographic variables summed to 101%. They presented a fake scale, the Acceptance of Sexual Minorities (ASM) Scale. They claimed that a random sample

of only 200 persons found exactly 100 each of right and left-handed mothers with LGB children. They claimed 85% of their sample was White compared to 96% of Singapore being Asian. In several instances, they reported different coefficients in the narrative versus the tables. One reported significance level was .2S5 (letters are not used in reporting significance levels). Their description of the ASM scale varied from a 5- to 7-point Likert scale. Most of their 12 to 15 (depending on the version) references were fabricated, irrelevant, or altered in some way. They claimed it should have been easy to detect that the paper was a fake (Rabak and Lan, 2024).

However, if one were to assume scholars are being truthful with their papers, that might not have been as easy as they believe. At the same time, there are even more reasons to doubt their paper's validity, as well as ways to statistically test its validity. Rabak and Lan (2024) specifically described the *Journal of Education, Society, and Behavioural Science* as "a fake journal", claiming that it had "all the hallmarks of a fake journal" because "They accept almost any kind of research in any field that's vaguely related to their journal name – even those that are not original or showed no results. The second reason we know it's a fake journal is because we wrote and submitted a fake scientific paper to them, and it got published." Furthermore, Rabak and Lan (2024) extensively criticized the research of Sullins (2015) and other conservatives, their fabricated paper being one part of that process of trying to discredit Sullins (2015). In particular, it should be noted that Rabak and Lan (2024) claimed that "Donald Paul Sullins' work is junk science" and that "Donald Paul Sullins' papers were published in a fake journal" – hence it is reasonable to conclude that a primary objective of trying to test the

journal was to discredit, not only the journal's review process, but moreover the research of Dr. Sullins, a professor of sociology from the Catholic University of America, even though his article had been published eight years earlier under a different journal editor, different journal title, and different reviewers. After the authors admitted to fabricating the article, it was removed from the journal's website.

Objectives

First, it may be useful for readers to consider the nature of the peer review process in general in academia, with particular attention to the question of whether top tier social science journals are immune to publishing papers with fabricated/fake data or results. Second, it may also be useful to illustrate possible explanations of how the editor and the three reviewers did not detect the fabricated nature of Rabak and Lan (2023) and, furthermore, how easy explanations might have excused the observed anomalies and still allowed for publication of this fake paper. Third, there are several ways in which scholars could have done further tests, on the Rabak and Lan (2023) paper (or other papers) to evaluate their scientific validity. Fourth, since the ultimate objective of Rabak and Lan (2023) was to discredit the journal and an article previously published in it (Sullins, 2015), it is appropriate to review the credibility of Sullins' (2015) article, even though his article had been published eight years earlier under a different journal name, editor, and different reviewers. Finally, implications and recommendations for peer review will be discussed.

Review of Literature on Peer Review in General and Rabak and Lan (2023) as an Example

Peer Review Processes in General

I think it is optimistic to think that all scholars, including undergraduate and graduate students, understand many of the details behind the process of peer review. Peer review has a long history, dating back thousands of years to ancient Greece (Shanmugam, 2022). Shanmugam (2022) reports that the first U.S. journal to adopt peer review was in 1818– *The American Journal of Science* at Yale University, edited and financially supported by Professor Benjamin Silliman. Peer review has strong advocates and strong critics, a polarized situation (Tennant & Ross-Hellauer, 2020, p. 2). While “peer review is routinely and widely criticized” (Tennant & Ross-Hellauer, 2020, p. 1), some even say that “the current system of blind peer review is obsolete” (Shanmugam, 2022, p. 21); yet Kassirer and Campion (1994) argue that peer review is indispensable.

In contrast, Shanmugam asserts that “the system of peer review is biased, unjust, unaccountable, incomplete, easily fixed, often insulting, usually ignorant, occasionally foolish, and frequently wrong” (p. 2). One problem may be that, as Radun (2023) has observed, that “recruiting peer reviewers has been increasingly difficult in recent years... so it is possible that even the best journals sometimes adopt the take-what-you-can-get strategy” (p. 335).

Furthermore, Shanmugam argues that “scientific progress is often made by departing from conventional wisdom,” yet conventional wisdom “often dictates what is being published by major scientific journals today,” thus having “negative effects on the peer-review process and on scientific progress” (p. 11). Peer reviewers may be “so tied to the conventional wisdom that they feel duty-bound to go to

extraordinary measures to find reasons for rejecting a manuscript with unconventional ideas (Shanmugam, 1986, 2022, p. 11). A possible result of this “conventional wisdom” problem is the fact that numerous very innovative papers that were initially rejected by scientific journals were later awarded Nobel prizes for their scientific advances (Campanario, 2009). In the area of social sciences, conventional wisdom may be highly correlated with so-called progressive political correctness.

How Did Rabak and Lan Survive Peer Review?

One could argue that any reasons given for the acceptance of a fabricated paper would be little more than “lame excuses.” Justification theory (Schumm, 2024) would support a tendency for editors and reviewers in such a situation to defend and try to explain their actions. However, a discussion of even weak/lame explanations may help future editors and reviewers from repeating past mistakes. First, my experience as an editor (for the journal *Marriage & Family Review*, 2010-2020) as well as peer reviewer for dozens of other journals, is that many submitted papers contain glaring errors of punctuation, spelling, grammar, and even typographical errors within their results. Sometimes there are English language deficiencies in general. If an editor or peer reviewer were to reject any paper that had such errors, they might have few to review, much less publish. Worse, the ability of many scholars to provide complete references, accurate and in consistent journal style, often seems lacking. Some journals have even given up the attempt to deal with this for new submissions, allowing authors to repair their references and make sure there is a 1:1 ratio of citations to references *after* the paper has been accepted. Some journals also will repair English

language issues in a paper *after* it has been accepted.

Thus, peer reviewers may tend to focus less on correcting references or weak English, even internal inconsistencies, than otherwise might be hoped. If you assume scholars are being honest and that journals will allow authors to correct their references after the paper is accepted, then it might not be as surprising that peer reviewers would overlook errors among references or language. It is also possible that the editor for the Rabak and Lan (2023) paper did not follow the journal’s own guidelines for emergency review (e.g., guidelines that include using more reviewers than normal and involving a second or third editor). Furthermore, it is widely recognized that detection of fraud is difficult even during multiple peer reviews (Radun, 2023, p. 332; Tennant & Ross-Hellauer, 2020, p. 3). Tennant and Ross-Hellauer went on to state that “fraudulent behaviour, or questionable research practices, still enter the scholarly record at high proportions, even though peer review occurs” (p. 7). Shanmugam (2022) asserts that peer review is “almost useless for detecting fraud” (p. 2).

When I received Rabak and Lan (2023) to review, I (mistakenly) operated from an assumption that the editor had already reviewed this paper as a possible scam and had cleared it. Did the editor notice any discrepancies between their names, university, and their source data (email or physical addresses)? An email of jinrabak@gmail.com was used to the editor). For example, when I submit a paper to a journal, my email is schumm@ksu.edu, which provides my last name, education source, and identifies a specific university while my physical address would tie me to a specific department or at least to the same town as Kansas State University. For a genuine university researcher to use a Gmail

address would be a surprise. I checked the July to December (2022) issues of *Marriage & Family Review* (under a new editor) and found only 20% (3/15) of corresponding authors used Gmail addresses.

Data or results fabrication is one of the most serious types of scientific misconduct, and it risks serious academic consequences (Nurunnabi & Hossain, 2019). Because of the presumed rarity of such fabrication, as well as its serious academic consequences, many editors and peer reviewers may not be inclined to suspect it. Indeed, I was not looking for signs of fabrication during my first review, although there was a potential question raised in my review regarding a lack of data for the ages of the women in the study, which made it difficult to check that data for validity or possible fabrication.

Rather, given the short deadline, I was looking for plausibility and being left-handed myself as a child, it made some sense that a left-handed mother might be more understanding of a sexual minority child's situation, even though the narrowness of the hypothesis was an issue in my mind—the expenditure of a lot of effort to collect such a large sample for such a narrow theory. Rabak and Lan (2024) seem to expect reviewers to be geographical experts, which I am not when it comes to the geographical sizes of cities I have never visited. While the authors (2023) claim they said they used a random sample, they actually said they used random household visits, which left open that they randomly visited households but only collected data from eligible participants, leading to the 200 cases they used. They never used the term “random sample” per se in Rabak and Lan (2023). While reviewers could be faulted for not noticing that Singapore is 96% Asian, the other reviewers (e.g., India, Hong Kong) lived much closer to Singapore than myself. Another subtle issue with Rabak and Lan

(2023) was that the longest sections of their paper were for the introduction and conclusion, when methodology usually occupies a much larger part of a paper in order to permit others to replicate the research.

As far as the reference issues, I trust authors to be truthful about their sources; generally, I look for issues of style deviation as much as anything else. Not expecting any scholar to fake references, I did not look for that as an issue and overlooked the problems there. Furthermore, my copy of the initial paper for Rabak and Lan (2023) did not contain 14 or 15 references as Rabak and Lan (2024) stated but rather only 12, which should have been a signal by itself, noted in my review—many journals will not accept papers that have fewer than 30 references. Regarding the numerical discrepancies, my approach is often to focus first on larger issues and on second review fine tune the process by focusing on whether numbers match up across narrative and tables. However, as best as I can tell from the website and my recollection, I did not get a chance to perform a second review, which normally allows a reviewer to verify if the recommendations in their first review have been considered by the paper's authors.

Finally, in their revision, the authors did not fully respond to my criticisms (e.g., did not provide estimates of the reliability of their measures in their current study, did not provide the range of ages of their participants), which I could never deal with since I did not get to see their revised version, as best I recall, at which point I could have noted that. One major issue for reviewers is when authors fail to take recommended changes, corrections, or improvements into account and ignore the advice given by reviewers, that becomes a red flag in itself. Had the review process gone to a second review, it is likely that

more of the issues would have been noted and perhaps the scam nature of the paper would have been detected.

On the other hand, the numbers within Tables 2, 3, and 4 did appear to be internally consistent, even though they differed from the narrative. When I write a paper, I usually do the tables first and then the narrative, leading me to focus on the results in tables first. The paper also presented an apparently significant interaction effect with a clear figure of the interaction pattern, a type of result often overlooked in many papers of lower quality. The large sample size, assuming it was valid, was also a strong plus for the paper since many studies prior have featured far smaller sample sizes for the relevant groups (e.g. Regnerus, 2012a,b; Sullins, 2015). The paper also appeared to address an issue that may not have been studied very much previously, a favorable situation in my opinion. Rejecting the paper simply because of its topic of research (e.g., LGBT research) or because it needed some improvements wasn't a reasonable or legitimate option in my opinion.

Methods for Detecting Fabricated Research

It is not clear what methods were used to generate the research results fabricated by Rabak and Lan (2023). We may never know *how* they did it unless they inform us themselves, which they did not do when they reported the fact of fabrication (but not the *how*) (Rabak and Lan, 2024). Perhaps others can speculate further on the *how* question.

Fortunately, *how* they did it is *not* a prerequisite for being able to assess whether

their reported results featured serious anomalies indicative of fabricated research. Here two general methods will be used. First, the ways in which Rabak and Lan (2023) attempted to design their fabricated paper will be considered. Were their methods sophisticated enough that they could not be excused merely as typographical errors? Second, are there ways in which deeper methods of analysis can be used to detect fabrication of results, methods that could be applied to the results reported in Rabak and Lan (2023)?

Results

Typographical Errors?

It is much more difficult to prove fraud than to speculate that it might have occurred. Let us suppose that a reviewer had gotten suspicious about Rabak and Lan (2023) and had noted the many apparent internal inconsistencies in their report. Would Rabak and Lan have just sent up a white flag and admitted their fraud? Justification theory (Schumm, 2024) would suggest not. By that theory, if you are caught doing something wrong, you have several possible responses: (1) Admit it and apologize, make amends; (2) Deny the situation occurred; (3) Deny responsibility even if it did occur; (4) Deny guilt even if responsible; (5) Blame someone else; (6) Justify what you did as if it were the right thing to do, for a greater cause; and if all else fails, (7) Minimize the level of the problem/mistake (no big deal, right?). How would this have worked out if a second review had found some or all the issues? Table 1 shows the issue and the likely response if noticed by a reviewer.

Table 1.

Had Rabak and Lan (2023) Been Challenged for Their Anomalies, How They Might Have Explained or Justified Each of Those Anomalies to Avoid Rejection of Their Fabricated Paper

Issue	Likely Response	Remarks
Author names were fake	We were trying to ensure a truly blind review	Some journals won't do blind reviews, so maybe they do feel this is needed
Author's university was fake	We were trying to ensure a truly blind review	Perhaps they don't want their university to find out if their article had been rejected
Area of Singapore was incorrect	Typographical error	Our typist messed it up
Demographic variable summed to 101%	Rounding error or typographical error	Actually 101% does occur in legitimate research
Acceptance of Minorities Scale was fabricated	We got the name of the scale wrong or we invented it ourselves	Our typist messed up again
Age of mothers, as reported, did not include decimals (e.g., average age was 43 rather than say, 43.2)	We inadvertently left off decimals three times for mothers' ages in Table 1	We will include those in our revision
The overall percentage for Blacks in Table 1 was 5 but was 5 for left-handed mothers and 6 for right-handed mothers; the overall percentage, given that the two groups are the same size, would have had to been 5.5 rather than 5.0	We made a typographical error for the overall percentage.	We will change that in our revision
Random sample?	We did not say we had a random sample per se, we had a random sample of households from which we picked eligible respondents	They never did say their sample was entirely random, only their initial selection of households
Exactly 200 participants, 100 each in each of two groups	We sampled each group until we got 100 in each	Researchers do this all the time, obtaining a quota for each group
85% of the sample was White	We meant to say 85% Asian and 5% White, sorry	That typist issue again
Different results in narrative versus the tables	The tables were correct but we did have typographical errors in the narrative	Typist issue

Significance level of 2S5	Sorry, another typographical error, easily corrected	Typist issue
Is the ASM 5-points per item or 7 points?	We meant five points, another typographical error	If 7 points, easier to show actual fraud, than if 5 points
References had issues	We were cutting and pasting references from other sources and may have made some mistakes; we will fix the problem before the paper is re-reviewed	This type of error does happen, especially when cutting and pasting from several sources
Article rejected for too many apparent errors	It is apparent that your journal, which published Sullins (2015) article, is biased and anti-gay	If the paper had been rejected, the authors could have easily accused the journal, its editor, and all three of its peer reviewers, including Dr. Schumm, of anti-LGBTQ+ bias as the underlying reason, regardless of the validity of the reviewers' comments

The point of Table 1 is that even if the authors had been caught with these issues, there would have been an easy, seemingly cogent response for each of them. If they had responded to each item in their reply to the editor, most editors would have taken them and their changes at face value and accepted the revised paper. And why not? So far there would have been no proof of fraud. Rabak and Lan (2024) seem to imply that if some reviewer had detected their anomalies, they would have admitted to faking the paper; but I would counter that with all they had at stake, I don't think they would have sent up a white flag yet. I think they would have worked around the issues as shown in Table 1 and still tried to get their fake paper published. Being an LGBTQ+ issue paper, I think an editor—not wanting to seem biased or homophobic—would have needed strong proof to reject the paper. Had the paper been rejected

outright, I would guess the next step would have been to accuse the journal, its editor, and the three reviewers of anti-LGBTQ+ bias, which would fit a narrative of the journal having accepted the Sullins (2015) because of that same presumed bias. Rabak and Lan (2024) may have seen the situation as a catch-22 “win, win”—rejection would imply bias, acceptance would imply incompetence. Either way, they win and have something with which to discredit the journal and hence, Sullins (2015).

Deeper Methods for Detecting Fabrication of Results; Further Tests for Rabak and Lan (2023)

Yet can there be strong proof that the paper was fabricated, proof beyond what could be explained away as a “few” typographical errors? That is what will be

discussed next—how such proof might have been provided.

Rabak and Lan (2024) seem to believe that it is rare for top tier journals to fail to identify fake research; nothing could be more incorrect. Elsewhere, we have explained how several scholars have succeeded (if one could call it that) at getting dozens of articles through peer review and published in top tier journals, despite having been fabricated (Schumm, Crawford, Lockett, AlRashed, & Ateeq, 2021, p. 29). For a prime example, top tier social science journals, for twenty or more years, failed to detect that a university distinguished professor from Florida State University had been submitting fabricated data and analyses without detection through multiple levels of peer review, including papers presented at conferences (which require peer review and then allow for audience review), having multiple co-authors (who presumably reviewed their own papers), journal peer review by two or more qualified experts, as well as options for post-publication peer review (Brezna, 2021; Pickett, 2020; Savolainen, 2024; Schumm, Crawford, Lockett, Ateeq, & AlRashed, 2023; Schumm, Crawford, Lockett, AlRashed, & Ateeq, 2024).

Thus, it is, without question, possible for top tier journals to fail to detect fabricated papers, not just once, but many times, and not just for a few months (Rabak and Lan, 2024), but for decades. It might even be the case that top tier journals are more likely to publish fake papers because of the higher status associated with publishing in such presumably credible journals; if a scholar can get fabricated research published in top tier journals, why waste one's time publishing in lower tier journals? Regardless, we have to agree with Rabak and Lan (2024) and others (Radun, 2023; Shanmugam, 2022; Tennant & Ross-Hellauer, 2020) that scientific peer review

can and perhaps even often does fail to detect issues with fraudulent data or results in scientific journals (Schumm, Crawford, Lockett, AlRashed, & Ateeq, 2021, p. 36).

There are at least nine ways to detect major research errors in published articles that have used small samples (Schumm, Crawford, Lockett, AlRashed, & Ateeq, 2021). Doing a deeper dive into the fake article, readers should have noticed that the range of the ASM scale with ten items had to run between 5 and 50 (5 point version) or between 7 and 70 (7 point version). None of the means presented were equal to or higher than 5.0, thus of nine mean scores reported, none were possible, unless the reported scores were for an average of the ten items in the ASM scale. Furthermore, they did not report actual ranges for their scale or for mothers' ages, which meant that our method of comparing standard deviations with ranges could not be used in its most ideal form. However, two of their standard deviations were the same, something Hartgerink and Wicherts (2016) said would be "extremely rare" (p. 6). Furthermore, the ratio of ideal range ($5 - 1 = 4$) to the standard deviations was between 3.92 and 6.67, within the expected range of 3 to 8 but still consistently on the high side. If one were to assume that the ideal range was 6 ($7 - 1 = 6$), then the ratio would have been between 5.88 and 10.00, with five of the 12 ratios greater than 8, outside the range of what is likely for the ratio of range divided by standard deviation. Applying the GRIM (Brown & Heathers, 2017) and/or SPRITE tests (Schumm, Crawford, & Lockett, 2019a, b) test to each of the nine means led to acceptable solutions for three of the nine means (3.52, 2.76, 2.33), "close" but not acceptable solutions for three of the nine (4.28, 4.44, 2.73), while three were impossible (4.52, 3.13, 2.92). The GRIM test can be checked by multiplying the mean by the sample size when a scale has only

integer values; the result should be an integer; for the last three “impossible” mean scores, that result did not yield integer values.

Furthermore, using the t-test calculator from Graphpad [[https://www.graphpad.com/quckcalcs/ttest 2/](https://www.graphpad.com/quckcalcs/ttest2/)] the four t-test values reported were incorrect, usually by substantial margins (13.13 rather than 7.52; 17.76 rather than 10.81; 10.45 rather than 3.38; 1.03 rather than 1.15). For two of the t-tests in Table 4, the degrees of freedom reported were incorrect (85 rather than 98; 111 rather than 98), which should have been noticed since the sum of the cases for those two t-tests (87 and 113) were not equal to the implied sample size of 100 (for the reported $df=98$). The chi-square statistic was incorrectly reported as 64.11, while it should have been 42.75; Cramer’s V was reported as .56 when it would have been .46.

While it is not clear that Benford’s Law applies to means and standard deviations (Schumm et al., 2023), in Tables 2, 3, and 4 of Rabak and Lan (2023), one would expect

about a third of the first 24 digits (i.e., 8) to be ones but only a single digit of one was reported. In terms of the last digits of the means and standard deviations in Tables 2, 3, and 4, there are the following counts for digits 0 through 9: 1, 1, 8, 6, 3, 0, 1, 1, 2, 1. Since the last digit should be approximately random (Schumm, 2023a,b), we should obtain roughly equivalent counts for each number, but instead, a one-sample chi-square test yields a value of 16.0 ($p < .03$), which indicates that the assignment of last digits was not random. Thus, if suspicions had been aroused about Rabak and Lan (2023), there were multiple ways to examine its validity statistically (Schumm, Crawford, Lockett, AlRashed, & Ateeq, 2021) as well as using freely available web-based statistical packages (Schumm, Dugan, Nauman, Sack, Maldonado, Conyac, & Patterson, 2021). Because these issues represent more than just a superficial analysis of Rabak and Lan (2023), it would have been much more difficult for them to refute or explain them away.

Table 2.

Additional Anomalies Found in Rabak and Lan (2023) That Would Have Been More Difficult to Explain/Justify After Review

Issue	Testing	Remarks
Range of ASM Scale	From scores of 5 to 50 or 7 to 70, outside of the ranges of reported scores	Rabak and Lan (2023) likely were reporting item averages, dividing scale scores by ten, the number of items
Two standard deviations were the same	Extremely rare occurrence	Sometimes rare occurrences do occur
If ASM Scale used five point items	Ratio of range to SDs would fall between 3.92 and 6.67, somewhat on the high side of expected range between 3 and 8	
If ASM scale used seven point items	Ratio of range to SDs would fall between 5.88 and 10.00, with 5/12 of	Most such ratios fall between 3 and 8, seldom above 8

	those reported greater than 8	
Mean scores and GRIM testing	Of the nine mean scores reported, six are impossible	Although three of the six are “close” to viable
Four t-test results and their degrees of freedom	All of the four test results reported are incorrect by wide margins using the data presented and half of the reported degrees of freedom are incorrect	Such major errors would be unlikely, even though one could claim typographical errors as the cause
Chi-square test and value of Cramer’s V	The chi-square test is incorrect as is Cramer’s V for the raw data reported	Maybe typographical errors could be blamed but unlikely a valid reason
Benford’s Law	Of the 24 values of means and SDs, eight would have been expected to start with a digit of “1” but only one of the eight did	While Benford’s Law does seem mainly to apply to regression coefficients, it may not always apply to means and SDs
Terminal digits as random	Terminal decimal point digits should be approximately random; the distribution of such digits was not random, with a chi-square of 16.00, $p < .03$.	

Criticism of Sullins (2015)

However, the point of submitting a scam paper to this journal was to try to discredit the journal, as an intermediate pathway to discrediting an earlier publication in the same journal by Dr. Paul Sullins (2015). First, that article was reviewed and edited by different scholars than used for Rabak and Lan (2023); thus, the possible “sins” of the editor and reviewers for Rabak and Lan (2023) were not relevant for the review process for Sullins (2015). Rabak and Lan (2024) proceed to present several criticisms of Sullins (2015), including that the gay or lesbian respondents could not have been married in 1995 when the survey had been conducted, that the measure of depression was broken into smaller subscales, that some of Sullins’s results favored the children of same-sex parents as if Sullins

did not mention that finding, that Sullins used three first person statements, that he cited Regnerus’s (2012a, b) study, and that Sullins reported results that were not statistically significant or were not tested for significance at all. Rabak and Lan (2024) concluded that Sullins (2015) was a “convoluted mess of a paper” that was “junk science at its highest form, and a low blow to the scientific community at large,” which may sound like psychological projection, or just plain hypocrisy, to those who dislike fabrication of research regardless of the rationale (i.e., that the ends justify the means).

Comments on Sullins (2015)

First, Sullins (2015) attempted to replicate earlier work by Wainright et al. (2004, 2006, 2008) using the same

variables, outcome measures, and sample. In terms of the marital status issue (how could same-sex parents be legally married before 2015?) respondents to the survey could report that they were in a marriage-like relationship as well as a legal marriage. 40% of the gay or lesbian parent couples reported that they were in a marriage-like relationship in the data used by Wainright et al. (see Sullins, 2015, p. 9). If Sullins was incorrect to credit those couples with a “married” status, one would have to agree that Wainright et al. (2004, 2006, 2008) were likewise incorrect. Sullins did use subscales from the CES-D but did so to replicate the methods used by the earlier analyses. If using subscales or items from the CES-D was an issue, it was an issue for both Sullins (2015) and Wainright et al. (2004, 2006, 2008). Sullins did mention results that were significant or not and, in contrast to the claims of Rabak and Lan (2023), mentioned that some results significantly favored the children of same-sex parents. Furthermore, in Sullins (2015), Figures 1 through 6 focused on interaction effects rather than main effects, but Sullins did note the presence of significant main effects that favored the children of same-sex parents within those six figures.

Sullins (2015) did use a first-person reference at three points; while this was frowned upon when I began my career as a professor in 1979, over time, with the support of feminist scholars, the trend has been more toward the use of personal pronouns. Sullins (2015) did cite the Regnerus study, which did receive much commentary afterwards, both more favorable (Regnerus, 2012b; Schumm, 2012) and unfavorable (reviewed in Schumm, 2015). Notably, some scholars agreed with Regnerus’s results inasmuch as relationship instability was found to mediate, even possibly explain, any associations between family structure and

various outcome measures (Allen & Price, 2020; Gates, 2015; Rosenfeld, 2015, Schumm, 2018, pp. 84-89). Rabak and Lan (2023) did not mention that Sullins (2015) had detected, in contrast to previous peer reviews of the same data, that only 17 of the same-sex female couples were lesbian parent couples of the 44 “same-sex” couples analyzed by Wainright et al. (2004), Wainright and Patterson (2006, 2008), and Patterson and Wainright (2012). Sullins (2015) also added data from three gay parent couples to his analyses, for a total of 20 genuine same-sex parent couples. Sullins (2015, pp. 11, 14-15) also was the first to notice that over 70% of the children of same-sex married couples reported having been forced to have sex against their will, of which a majority had been molested by a family caregiver (not necessarily a same-sex parent) in their past.

That issue is plausible—for example, Balsam, Rothblum, and Beauchaine (2005) found that nearly 44% of lesbians and 48% of bisexual women in their study had reported childhood sexual abuse. For all their criticisms, some justified, of conservatives, Rabak and Lan (2024) did not mention parallel criticisms of progressive scholars which, in contrast to their 2024 report, had been published in peer reviewed journals (Schumm, 2015, 2021; Schumm, 2020a, b; Schumm & Crawford, 2018; 2019a, b; 2020a, b; 2021; Schumm, Pakaluk, & Crawford, 2020), along with an article that commented on issues in both conservative and progressive scholars’ research (Schumm & Crawford, 2023), including that of Dr. Sullins. Schumm’s research has been subjected to analysis and criticism as well, by conservative scholars (Sutton & Cretella, 2018).

Discussion

Implications

The article by Rabak and Lan (2023) is not alone in containing a variety of anomalies that should lead to questions from reviewers. I have recently reviewed more than one paper where the numbers in the narrative did not match the numbers reported in the tables. Many papers lack adequate referencing and citations. Some articles (Elwood et al., 2020) have referenced as keystones of their literature review, articles (Hatzenbuehler, 2014) that have been retracted; specifically, Elwood et al. cited Hatzenbuehler et al. (2014) stating that “Furthermore, lesbians, gay men, and other sexual minorities who live in American regions with greater levels of prejudice are more likely to die approximately 12 years prematurely from cardiovascular diseases, homicide, suicide, and other causes than sexual minorities who live in more tolerant areas” (p. 935).

Furthermore, some articles (Elwood et al., 2020, p. 939) claim to represent lesbians and gay men as sexual “minorities” but those particular minorities actually scored higher on income levels and educational levels than did all respondents. Some papers have reported results as non-significant that when re-analyzed yielded results that were significant—and vice versa (for examples, see Schumm and Crawford, 2024). Some papers had results with small effect sizes that were significant but had results with larger effect sizes that were not significant (see Schumm and Crawford, 2023, p. 202-203). I’ve seen cases where the results were significant but by dividing the sample into smaller subgroups, the results appeared to become non-significant (Schumm & Crawford, 2023, 2024). However, outright fabrication of results is a far more serious academic problem than such issues, as relevant as they may be for assessing the

validity of research reports based on authentic data and non-fabricated results.

Furthermore, if a scholar wants to slant their research to either find or not find significant results, there are dozens of ways to do this subtly in ways that many reviewers may not notice (Schumm & Crawford, 2024). This problem of confirmation bias (Schumm, 2015, 2021) appears to be widespread. As long as desired or appealing results are obtained, research that is eventually proven to be largely fraudulent can be accepted for years, even decades, without any suspicion or complaints, despite multiple levels of peer review (e.g., informal faculty meetings to discuss ongoing research projects, conferences, papers submitted to academic journals papers after review by multiple co-authors, post-publication critiques). Reviewers need to be alert for such issues.

Finally, since the ultimate objective of the fabricated, fake Rabak and Lan (2023) paper appears to have been to discredit the journal and a particular article previously published in it (Sullins, 2015), it seems appropriate to consider the issue of conflicts of interest (Radun, 2023). By addressing Dr. Sullins’ research in an ad hominem manner, Rabak and Lan (2024) exposed their own conflict of interest. Post-publication reviews can be helpful in advancing social science, but such reviews should be even-handed and represent both the positives and negatives of articles under review (Lanier, 2021). Radun concludes that conflicts of interest among editors and peer reviewers that lead to “biased review[s] with personal insults” (p. 338), “especially regarding controversial issues in small research fields, [are] indeed dangerous and harmful to the scholarly community” (p. 339). The implication, that Radun (2023) has noted, is that “if such [peer] reviews contain personal insults or question the authors’ motivation, they

should always be discarded because they are clear signs of a serious COI [conflict of interest]” (p. 334). Reviewers should be alert for such issues.

Recommendations

I believe that today graduate students at some universities are receiving less training in how to do a peer review of a journal article. Learning how to do peer review is not a “sexy” kind of training. It involves careful attention to detail and takes time. If you start with the assumption that papers will not have inconsistencies or errors because they are by scholars with Ph.D.’s, you may not even look for such things and, not looking for them, overlook them, that is, fail to “see” or detect them. Hence, it is my recommendation that graduate programs include academic and practical training on peer review.

The editor of the *Journal of Education, Society, and Behavioural Science* should maintain the article on the journal’s website so comments about it can be fact checked and so it can be used for training others in the detection of fraudulent research. However, the word FABRICATED and RETRACTED in large letters should be placed across the first page of the article. Journals that provide for emergency review must be especially careful to guard against authors trying to take advantage of their emergency procedures. Editors must assume at least partial responsibility for detecting and, if possible, desk rejecting scam papers because the editor is the primary and initial interface between the paper’s author and the peer review process, even if editors depend on peer reviewers to identify or prove fraudulent work.

Graduate programs should clarify that criticism of research should not be received as disrespect or emotional harm; without criticism, research is apt to degenerate in

quality and value. Constructive criticism can and should be used to lead to higher quality research. At the same time, due to constraints of time and funding, all research will feature limitations, including even simple typographical errors; thus, criticism cannot mean that a research report has no value—it just may have limitations that may not be correctable. In some situations, internal inconsistencies such as those in Rabak and Lan (2023) may reflect simple typographical errors, nothing more. Limitations are important because they warn us to limit the applicability of any given bit of research, no matter how interesting or in tune with our own biases.

Editors and peer reviewers should know that faked research can be detected with a variety of methods; even if they are not familiar with those methods, if they suspect a research report has been faked, they should contact scholars who can confirm or disconfirm their concerns. Editors and peer reviewers should also be alert for conflicts of interest as indicated by attempts to discredit persons rather than merely offering criticism of specific aspects of the research reports of other scholars. Some journals have begun a process of sending out papers before peer review to an independent scholar for an assessment of whether the paper is deserving of being sent out for peer review, likely an attempt to, among other things, detect serious problems such as fabricated research prior to peer review and, perhaps, to reduce the need for further reviews and the overall burden on the journal’s support team.

Conclusion

The fact that any particular journal fails to detect any one fake or fabricated paper in the process of peer review does not demonstrate that the journal is lower tier or lacks credibility. Numerous high quality scholarly

journals have failed to detect a large number of faked papers. One must ask if it is worth engaging in one of the most serious forms of scientific misconduct in order to try to discredit a journal, much less to discredit an author in that journal several years prior, especially when the editors and reviewers were different? If a person is willing to lie about science in order to criticize someone, should they be believed about their other claims? If a person feels it is acceptable to lie for the “greater good,” why not lie about lesser things as well? In this particular case, not only did the Rabak and Lan (2023) misrepresent themselves but they appealed to the emotional desires of editors and peer reviewers to help them out, presenting themselves, at least to the journal’s editor, as a female graduate student with a need for an extraordinarily fast review, not only deceiving but even betraying a remarkable kindness from everyone involved in the expedited peer review process.

However, improvements in journal operations by editors and peer reviewers can increase the chances of detecting anomalies in submitted papers and analyzing them in sufficient depth to detect likely situations of data or results fabrication. If papers seem to present conflicts of interest in terms of *ad hominem* attacks on other scholars, further investigation may be needed, and such papers should be revised without such attacks before publication. While any paper may feature typographical errors or missing references, patterns of multiple such errors may indicate serious problems, even potential fraud, and should lead to very careful peer and editorial review.

References

- Allen, D., & Price, J. (2020). Stability rates of same-sex couples: With and without children. *Marriage & Family Review*, 56, 51-71.
- Balsam, K. F., Rothblum, E. D., & Beauchaine, T. P. (2005). Victimization over the life span: A comparison of lesbian, gay, bisexual, and heterosexual siblings. *Journal of Consulting and Clinical Psychology*, 73(3), 477-487. <https://doi.org/10.1080/01494929.2019.1630048>
- Breznau, N. (2021). Does sociology need open science? *Societies*, 11(1), 9-25. <https://doi.org/10.3390/soc1101009>
- Brown, N. J., & Heathers, J. A. (2017). The GRIM test: A simple technique detects numerous anomalies in the reporting of results in psychology. *Social Psychological and Personality Science*, 8(4), 363-369. <https://doi.org/10.1177/1948550616673876>
- Campanario, J. M. (2009). Rejecting and resisting Nobel class discoveries: Accounts by Nobel Laureates. *Scientometrics*, 81, 549-565. <https://doi.org/10.1007/s11192-008-2141-5>
- Elwood, W. N., Irvin, V. L., Liu, B., Lee, R., & Breen, N. (2020). Health-related influences of extending marital benefits to same-sex couples: Results from the California Health Interview Survey. *Family Relations*, 69(5), 934-943. <https://doi.org/10.1111/fare.12507>
- Gates, G. J. (2015). Marriage and family: LGBT individuals and same-sex couples. *The Future of Children*, 25(2), 67-87. <https://doi.org/10.1353/foc.2015.0013>
- Hartgerink, C. H. J., & Wicherts, J. M. (2016). Research practices and assessment of research misconduct.

- ScienceOpenResearch*.
<https://doi.org/10.14293/S2199-1006.1.SOR-SOCSCI.ARY5BI>, v1.
- Hatzenbuehler, M. L., Bellatorre, A., Lee, Y., Finch, B., Muennig, P., & Fiscella, K. (2014). Structural stigma and all-cause mortality in sexual minority populations. [Retracted] *Social Science and Medicine*, 103, 33-41. <https://doi.org/10.1016/j.socscimed.2019.112463>
- Hergovich, A., Schott, R., & Burger, C. (2010). Biased evaluation of abstracts depending on topic and conclusion: Further evidence of a confirmation bias within scientific psychology. *Current Psychology*, 29, 188-209. <https://doi.org/10.1007/s12144-010-9087-5>
- Johnson, V. E., Payne, R. D., Wang, T., Asher, A., & Mandal, S. (2017). On the reproducibility of psychological science. *Journal of the American Statistical Association*, 112(517), 1-10. <https://doi.org/10.1080/01621459.2016.1240079>
- Kassirer, J. P., & Champion, E. W. (1994). Peer review: Crude and understudied, but indispensable. *JAMA*, 272, 96-97. <https://doi.org/10.1001/jama.272.2.96>
- Lanier, W. L. (2021). Dealing with inappropriate-, low quality-, and other forms of challenging peer review, including hostile referees and inflammatory or confusing critiques: Prevention and treatment. *Accountability in Research*, 28(3), 162-185. <https://doi.org/10.1080/08989621.2020.1815010>
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, 2(2), 175-220. <https://doi.org/10.1037/1089-2680.2.2.175>
- Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., Fidler, F., Hilgard, J., Kline Struhl, M., Nuijten, M. B., & Rohrer, J. M. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology*, 73(1), 719-748. <https://doi.org/10.31234/osf.io/ksfvq>
- Nurunnabi, M., & Hossain, M. A. (2019). Data falsification and question on academic integrity. *Accountability in Research*, 26(2), 108-122. <https://doi.org/10.1080/08989621.2018.1564664>
- Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716. <https://doi.org/10.1126/science.aac4716>
- Patil, P., Peng, R. D., & Leek, J. T. (2016). What should researchers expect when they replicate studies? A statistical view of replicability in psychological science. *Perspectives on Psychological Science*, 11(4), 539-544. <https://doi.org/10.1177/174569161664636>
- Patterson, C. J., & Wainright, J. L. (2012). Adolescents with same-sex parents: Findings from the National Longitudinal Study of Adolescent Health. In D. M. Brodzinsky & A. Pertman (Eds.), *Adoption by lesbians and gay men: A new*

- dimension in family diversity*. New York: Oxford University Press.
- Pickett, J. T. (2020). The Stewart retractions: A quantitative and qualitative analysis. *Econ Journal Watch*, 17(1), 152-190.
- Rabak, J., & Lan, H. G. (2023). Left-handed mothers and LGBTQ+ child acceptance in Singapore: Exploring the link through early life rejection. *Journal of Education, Society, and Behavioural Science*, 36(4), 13-18. <https://doi.org/10.9734/jesbs/2023/v36i41216>
- Rabak, J., & Lan, H. G. (2024). *Anti-LGBTQ fake science is spreading in Singapore*. <https://heckinunicorn.com/blogs/heckin-unicorn-blog/anti-lgbtq-fake-science-is-spreading-in-singapore?> Updated January 8, 2024, 54 pages. Singapore, Republic of Singapore.
- Radun, I. (2023). Nonfinancial conflict of interest in peer review: Some notes for discussion. *Accountability in Research*, 30(6), 331-342. <https://doi.org/10.1080/08989621.2021.1989677>
- Regnerus, M. (2012a). How different are the adult children of parents who have same-sex relationships? Findings from the New Family Structures Study. *Social Science Research*, 41, 752-770. <https://doi.org/10.1016/j.ssresearch.2012.03.009>
- Regnerus, M. (2012b). Parental same-sex relationships, family instability, and subsequent life outcomes for adult children: Answering critics of the New Family Structures Study with additional analyses. *Social Science Research*, 41, 1367-1377.
- Rosenfeld, M. J. (2015). Revisiting the data from the New Family Structure Study: Taking family stability into account. *Sociological Science*, 2, 478-501. <https://doi.org/10.1016/j.ssresearch.2012.08.015>
- Savolainen, J. (2024). Unequal treatment under the law: Race, crime, and retractions. *Current Psychology*, 43(17), 16002-16014. <https://doi.org/10.1007/s12144-023-04739-2>
- Schumm, W. R. (2015). Navigating treacherous waters – one researcher’s 40 years of experience with controversial scientific research. *Comprehensive Psychology*, 4, 24, 1-40. <https://doi.org/10.2466/17.CP.4.24>
- Schumm, W. R. (2018). *Same-sex parenting research: A critical assessment*. London: Wilberforce Press.
- Schumm, W. R. (2020a). Changes over the decades in selected LGBTQ research findings. *JSM Sexual Medicine*, 4(2), 1029, 1-8. <https://doi.org/10.47739/2578-3718/1029>
- Schumm, W. R. (2020b). Avenues for future LGBT theory and research. *JSM Sexual Medicine*, 4(2), 1031, 1-7.
- Schumm, W. R. (2021). Confirmation bias and methodology in social science: An editorial. *Marriage & Family Review*, 57(4), 285-293. <https://doi.org/10.1080/01494929.2021.1872859>
- Schumm, W. R. (2023a). A further examination of excesses or deficits of terminal zeroes in scientific research: Using binomial testing to assess probabilities for different anomaly classifications as a function of different sample sizes of coefficients with terminal digits. *Biomedical Journal of Scientific & Technical Research*, 52(4), 43992-43995.

- <https://doi.org/10.26717/BJSTR.2023.52.008293>
- Schumm, W. R. (2023b). Assessing last decimal digits in retracted research articles: Accounting for missing zeroes. *SCIREA Journal of Sociology*, 7(4), 276-286. <https://doi.org/10.54647/sociology841134>
- Schumm, W. R. (2024). *Justification theory*. Unpublished working paper, Manhattan, Kansas.
- Schumm, W. R., & Crawford, D. W. (2018). How have other family journals compared in LGBT research to the top seven journals in the field? *Marriage & Family Review*, 54(6), 521-530. <https://doi.org/10.1080/01494929.2018.1460145>
- Schumm, W. R., & Crawford, D. W. (2019a). Evaluating the quality of literature reviews in the social sciences: Developing a measure of quality with an illustration. *Psychology Research & Applications*, 1(2), 47-69. <https://dx.doi.org/10.22606/pr.2019.12003>
- Schumm, W. R., & Crawford, D. W. (2019b). Scientific consensus on whether LGBTQ parents are more likely (or not) to have LGBTQ children: An analysis of 72 social science reviews of the literature published between 2001 and 2017. *International Journal of Women's Studies*, 20(7), 1-12. Available at: <https://vc.bridgew.edu/jiws/vol20/iss7/1>
- Schumm, W. R., & Crawford, D. W. (2020a). Assessing citation bias in scientific literature. *American Journal of Biomedical Science & Research*, 10(3), 274-276. <https://doi.org/10.34297/AJBSR.2020.10.001514>
- Schumm, W. R., & Crawford, D. W. (2020b). Is research on transgender children what it seems? Comments on recent research on transgender children with high levels of parental support. *The Linacre Quarterly*, 87(1), 9-24. <https://doi.org/10.1177/0024363919884799>
- Schumm, W. R., & Crawford, D. W. (2021). Can social desirability be detected statistically at the macro-sociological level? *Psychology Research & Applications*, 3(1), 14-28. <https://doi.org/10.22606/pr.2021.31002>
- Schumm, W. R., & Crawford, D. W. (2023). Difficulties with methodology in social science research with controversial issues regarding human sexuality. *The Linacre Quarterly*, 90(2), 194-216. <https://doi.org/10.1177/00243639221082213>
- Schumm, W. R., & Crawford, D. W. (2024). Sixteen approaches for deconstructing theory, research, or literature reviews claiming a “no difference” (null) result in the social sciences. *The Natural Family: A Journal of Research and Public Policy*, 36(3-4), 155-176.
- Schumm, W. R., Crawford, D. W., Higgins, M., Lockett, L., AlRashed, A., & Bin Ateeq, A. (2018). Estimating the standard deviation from the range: A replication of analyses of demographic data reported in *Marriage & Family Review* 2016 – 2017. *Marriage & Family Review*, 54(8), 777-792. <https://doi.org/10.1080/01494929.2018.1500128>

- Schumm, W. R., Pakaluk, C. R., & Crawford, D. W. (2020). Forty years of confirmation bias in the social sciences: Two cases of selective citations. *Internal Medicine Review*, 6(4), 1-14. <https://doi.org/10.18103/imr.v6i4.875>
- Schumm, W. R., Crawford, D. W., & Lockett, L. (2019a). Patterns of means and standard deviations with binary variables: A key to detecting fraudulent research. *Psychology Research and Applications*, 1(4), 112-118. <https://doi.org/10.26717/BJSTR.2019.23.003851>
- Schumm, W. R., Crawford, D. W., & Lockett, L. (2019b). Using statistics from binary variables to detect data anomalies, even possibly fraudulent research. *Biomedical Journal of Scientific and Technical Research*, 23(1), 17151-17153. <https://doi.org/10.22606/pr.2019.14004>
- Schumm, W. R., Crawford, D. W., Lockett, L., AlRashed, A., & bin Ateeq, A. (2021). Nine ways to detect possible scientific misconduct in research with small ($N < 200$) samples. *Psychology Research and Applications*, 3(2), 29-40. <https://doi.org/10.22606/pr.2021.32001>
- Schumm, W. R., Crawford, D. W., Lockett, L., AlRashed, A., & bin Ateeq, A. (2024). Research anomalies in criminology: How serious? How extensive over time? And who was responsible? *Accountability in Research*, online advance. <https://doi.org/https://doi.org/10.1080/08989621.2023.2241127>
- Schumm, W. R., Crawford, D. W., Lockett, L., bin Ateeq, A., & AlRashed, A. (2023). Can retracted social science articles be distinguished from non-retracted articles by some of the same authors, using Benford's Law or other statistical methods? *Publications*, 11, 14, 1-13. <https://doi.org/10.3390/publications11010014>
- Schumm, W. R., Dugan, M., Nauman, W., Sack, B., Maldonado, J., Conyac, C., & Patterson, C. (2021). Using free websites to perform statistical calculations in basic statistics courses at high school or college levels. *SunKrist Sociology and Research Journal*, 2(1), 1009, 1-7.
- Shanmugam, G. (1986). Conventional wisdom and scientific progress. *Geology*, 14(8), 718. [https://doi.org/10.1130/0091-7613\(1986\)14<718a:CWASP>2.0.CO;2](https://doi.org/10.1130/0091-7613(1986)14<718a:CWASP>2.0.CO;2)
- Shanmugam, G. (2022). The peer-review problem: A sedimentological perspective. *Journal of the Indian Association of Sedimentologists*, 39(3), 24. <https://doi.org/10.51710/jias.v39i1.243>
- Sullins, D. P. (2015). The unexpected harm of same-sex marriage: A critical re-appraisal, replication, and re-analysis of Wainright and Patterson's studies of adolescents with same-sex parents. *British Journal of Education, Society, & Behavioural Science*, 11(2), 1-22. <http://dx.doi.org/10.2139/ssrn.2589129>
- Sutton, P. M., & Cretella, M. (2018). A review and summary of Walter Schumm's *Same-Sex Parenting Research: A Critical Assessment*. *Journal of Human Sexuality*, 9, 101-107.

- Tennant, J. P., & Ross-Hellauer, T. (2020). The limitations to our understanding of peer review. *Research Integrity and Peer Review*, 5: 6, 1-14. <https://doi.org/10.1186/s41073-020-00092-1>
- Wainright, J. L., Russell, S. T., & Patterson, C. J. (2004). Psychosocial adjustment, school outcomes, and romantic relationships of adolescents with same-sex parents. *Child Development*, 75(6), 1886-1998. <https://doi.org/10.1111/j.1467-8624.2004.00823.x>
- Wainright, J. L., & Patterson, C. J. (2006). Delinquency, victimization, and substance use among adolescents with female same-sex parents. *Journal of Family Psychology*, 20(3), 526-530. <https://doi.org/10.1037/0893-3200.20.3.526>
- Wainright, J. L., & Patterson, C. J. (2008). Peer relations among adolescents with female same-sex parents. *Developmental Psychology*, 44(1), 117-126. <https://doi.org/10.1037/0893-3200.20.3.526>