Fatal flaws in “The relationship between personality traits and marital satisfaction: a systematic review and meta-analysis”

Steven G. Ludeke*

*Corresponding author

University of Southern Denmark, Odense, Denmark

Steven Ludeke@gmail.com

Submitted to *Meta-Psychology*. Participate in open peer review by sending an email to open.peer.reviewer@gmail.com. The full editorial process of all articles under review at Meta-Psychology can be found following this link:

https://tinyurl.com/mp-submissions

You will find this preprint by searching for the first author's name.

Abstract:

The core results from a recent *BMC Psychology* meta-analysis represent errors of extraordinary magnitude and implausibility, with reported correlations between marital satisfaction and the Big Five traits as high as r = .90. Inspection of the paper reveals that standardized z-statistics were misinterpreted as correlation coefficients, producing effect sizes far beyond plausible limits. Despite the error's remarkable implausibility, the article has remained unretracted for five years and accumulated over two hundred citations, none expressing doubt. The handling journal acknowledges the problem but reports no process for publishing error-only reports. This episode illustrates a failure of science's self-correction mechanisms and highlights the demotivating effect of uncorrected errors on conscientious researchers. It also raises broader questions about journal accountability and impact-factor systems that equate venues regardless of their capacity for post-publication correction.

Keywords: error correction; personality; marital satisfaction

Discovery of an Impossible Result

While reading an unusually strong bachelor's thesis on personality and relationship quality, I was brought up short by an impossible result. The student cited a recent meta-analysis appearing in *BMC Psychology*, a journal from the Nature Springer group, which reported correlations between marital satisfaction and the Big Five traits of r = .90 (Conscientiousness), .86 (Agreeableness), .83 (Extraversion), .78 (Openness), and –.44 (Neuroticism; Sayehmiri et al., 2020).

I assumed this was a transcription error: Numbers this implausible are not unreasonably attributed to the student copying the wrong results down. Upon checking the source, I saw however that the student had transcribed the results exactly; the journal article itself repeatedly presented these values and discussed them as genuine findings.

Despite my previous work documenting serious quantitative errors in published research, this one struck me as extraordinary. The magnitude of the discrepancy between expectations (reviewed below) and reported meta-analytic results struck me as well beyond anything that could plausibly survive peer review—and yet the article was being cited more rapidly than many of the papers I find most important in the field. My reaction went well beyond surprise: I genuinely albeit temporarily raised my estimate that I might have lost contact with reality, and wrote to a colleague for a (quite literal) sanity check. A short refresher on why such results are so impossible that they provoked such a response is therefore warranted, as it provides crucial context for interpreting what this incident reveals about contemporary psychological research and publishing practices.

Commonly-Observed Personality Effect Sizes, and the Implications of More Pronounced Effects

As a rule, the most potent personality predictors of any given outcome typically yield correlations in the moderate range. (See, for example, studies of counterproductive work behaviors [Berry et al., 2007], academic performance [Poropat, 2009], and externalizing psychopathologies [Ruiz et al., 2008].) Meta-analytic correlations with personality traits rarely approach .50, and in the exceptional cases that do reach that range—e.g., Neuroticism and internalizing disorders (Ormel et al., 2013)—it's not uncommon for researchers to voice skepticism of the differentiability of the two constructs. Psychometric correlations that are particularly large (>.80) are widely recognized as indicating the traits under study are psychometrically redundant (Lawson & Robins, 2021). An early meta-analysis on the link between self-reported personality and marital satisfaction—precisely the topic explored by Sayehmiri et al. (2020)—gives no reason to expect this domain to be exceptional, reporting effects ranging from −.26 for Neuroticism to just .08 for Openness (Heller et al., 2004).

By contrast, the results reported by Sayehmiri et al. (2020)—one unexpectedly large effect (.44 with Neuroticism) and four extraordinary ones (ranging between .77 and .90 for the other Big Five traits)—would open intriguing questions well beyond marital satisfaction. They would suggest that, at least in the Iranian context of the studies meta-analyzed by Sayehmiri et al. (2020), most Big Five traits (excluding Neuroticism) are not only highly predictive of marital satisfaction but also nearly isomorphic with it—and with each other. This could represent fascinating evidence for the General Factor of Personality, a striking incidence of cultural differences, and a powerful rebuttal to Mischel and the situationists to boot.

Re-Examining the Evidence: The Mislabelled Z-Scores

Inspection of the original paper's results table provided an important clue to the unexpected results. Two adjacent columns were labelled "Z-score" and "R." When the values in the *Z* column are interpreted as correlation coefficients, they match almost perfectly the earlier meta-analytic results reported by Heller et al. (2004); the "R" column—the one the authors described multiple times in the text as correlations—deviates by roughly fifteen times as much. Table 1 presents these side-by-side comparisons.

Table 1. Comparison of meta-analytic results between the Big Five and marital satisfaction			
	Heller et al (2004)	Sayehmiri et al (2020)	
		"Z score"	"R"
Neuroticism	-0.26	-0.22	0.439
Extraversion	0.14	0.16	0.833
Agreeableness	0.24	0.14	0.855
Conscientiousness	0.22	0.26	0.9
Openness	0.08	-0.04	0.777
Note. The results reported from Heller et al. (2004) are from their "Avg. r" column in Table 5. Given that there is no indication in Sayehmiri et al. (2020) that they performed any corrections for measurement error, reliability, or measurement artifacts, the "Avg. r" results are presumably the appropriate comparisons rather than the "ρ" results from the same table (which are, in any event, only 10-20% larger).			

The simplest explanation is that standardized *z* statistics were mistaken for correlation coefficients at some point in the analytic workflow or manuscript preparation. The English-language primary studies included in the meta-analysis confirm the implausibility of the published numbers. Across three accessible papers (Amiri et al., 2011; Javanmard & Garegozlo, 2013; Sadeghi & Mozhdehi, 2016), the largest observed coefficient between any Big Five trait and marital satisfaction was .50, and twelve of fifteen were below |.32|. The notion that the true population effects could average above .80 across studies is not credible.

The error is categorical, not interpretive. Under no reasonable statistical understanding can the published results be valid. Given such a basic and extraordinary error, the analytic procedures and quality-control practices of the original authorship team behind the report warrant considerable independent re-examination by qualified analysts.

Widespread Uptake Without Scrutiny

The implausible correlations have not remained obscure. At the time of writing, *Google Scholar* listed roughly 240 citations to the article in the five years since its 2020 publication, including 32 added after the journal was first notified of the error. A spot-check of dozens of these citing papers found none expressing skepticism about the reported magnitudes. Several used the results substantively—re-analyzing or interpreting them as genuine findings—while many more repeated the coefficients uncritically in literature reviews or introductions. The error propagation is not limited in reach, having even affected articles published in elite outlets within the field (e.g., Esplin et al., 2025; Grau et al., 2025).

Attempts to Correct the Record

A paper briefly documenting the error was submitted to *BMC Psychology* on June 26, 2025. It was desk-rejected on July 14, appealed on July 30, and subsequently rerouted through several

administrative exchanges. Forms were misplaced, replaced, and resubmitted, with weeks of silence between replies. On September 26, the appeal was rejected, and I was advised to contact the handling editor directly.

Over the next several days, correspondence with that editor (Dr. Menglu Yang) confirmed that *BMC Psychology* "does not have a specific article type" for work whose sole purpose is to document an error. As of November 2025, the article remains online and unretracted. No timeline for correction has been communicated.

The correspondence gives no indication that the editor or publisher has evaluated the statistical substance of the error. The obstacle appears procedural, not substantive, revealing a gap in the publisher's ability to address clear-cut mistakes once an article is in print.

Incentives and Motivation

Within the replication crisis, discussions of incentives have focused mainly on those relevant to identifying and correcting flawed research. These do matter: detecting and documenting errors requires many hours of unfunded work, and few journals or universities recognize it. Other domains, such as software security, reward flaw detection through "bug-bounty" programs; science expects comparable diligence without compensation or even meaningful acknowledgment.

Less often considered, but perhaps more consequential, is the effect of bad science on researchers' intrinsic motivation for their own original empirical work. The social or scientific value of any well-conducted study depends on the reliability of the field it builds on. When researchers perceive that (a) much existing work is of low quality and (b) influence is weakly linked to quality, the expected social and scientific value of doing careful research declines. Because academic research depends heavily on intrinsically motivated individuals—those who remain in science despite limited extrinsic rewards—tolerating low-quality work risks eroding the motivation that sustains progress. Findings from research on "meaningful work," "callings," "moral injury," and "public-service motivation" all converge on this point: maintaining intrinsic motivation requires a culture of quality control and credible self-correction.

Lessons for Science's Claim of Self-Correction

Science's credibility rests on its capacity for self-correction. Yet this case shows an error so obvious that it should have been detected at any stage—peer review, editorial handling, post-publication citation—which nevertheless persisted for five years while quickly accumulating citations in the field. The episode is highly inconsistent with an optimistic view of contemporary psychological science, instead better aligning with the position that major overhauls to incentive structures and publication practices are urgently needed.

The episode also bears on narrower questions, such as the role of *BMC Psychology* within the field. If a journal cannot or will not correct known errors, its publications cannot serve as

equivalent academic currency to those that can. Journals lacking rigorous pre- and post-publication review processes arguably must be treated, in effect, as self-verified preprints. Impact factors that treat all journals alike obscure this difference. A credible alternative would be an impact-factor provider that withholds rankings from outlets that lack a formal channel for publishing credible error corrections to their own articles.

The episode began with a student whose only error was trusting a published paper, and a supervisor briefly wondering if he'd lost his mind. It ends with a community still largely acting as though nothing is amiss. In that sense, the "sanity check" remains incomplete—not for one researcher, but for the field itself.

References

Amiri, M., Farhoodi, F., Abdolvand, N., & Bidakhavidi, A. R. (2011). A study of the relationship between Big-five personality traits and communication styles with marital satisfaction of married students majoring in public universities of Tehran. *Procedia - Social and Behavioral Sciences*, *30*, 685-689. https://doi.org/10.1016/j.sbspro.2011.10.132

Berry, C. M., Ones, D. S., & Sackett, P. R. (2007). Interpersonal deviance, organizational deviance, and their common correlates: a review and meta-analysis. *J Appl Psychol*, *92*(2), 410-424. https://doi.org/10.1037/0021-9010.92.2.410

Esplin, C. R., Rasmussen, B. D., Hatch, S. G., Hawkins, A. J., & Braithwaite, S. R. (2025). Neuroticism and relationship quality: A meta-analytic review. *J Pers Soc Psychol*, *128*(3), 594-610. https://doi.org/10.1037/pspi0000463

Grau, I., Miketta, L., Ebbeler, C., & Banse, R. (2025). Cultural Differences in Correlations With Marital Satisfaction: The Vulnerability-Stress-Adaptation Model in 65 Countries. *Journal of Cross-Cultural Psychology*. https://doi.org/10.1177/00220221251323216

Heller, D., Watson, D., & Hies, R. (2004). The role of person versus situation in life satisfaction: a critical examination. *Psychol Bull*, *130*(4), 574-600. https://doi.org/10.1037/0033-2909.130.4.574

Javanmard, G. H., & Garegozlo, R. M. (2013). The Study of Relationship Between Marital Satisfaction and Personality Characteristics In Iranian Families. *Procedia - Social and Behavioral Sciences*, *84*, 396-399. https://doi.org/10.1016/j.sbspro.2013.06.573

Lawson, K. M., & Robins, R. W. (2021). Sibling Constructs: What Are They, Why Do They Matter, and How Should You Handle Them? *Pers Soc Psychol Rev*, *25*(4), 344-366. https://doi.org/10.1177/10888683211047101

Ormel, J., Jeronimus, B. F., Kotov, R., Riese, H., Bos, E. H., Hankin, B., Rosmalen, J. G. M., & Oldehinkel, A. J. (2013). Neuroticism and common mental disorders: meaning and utility of a complex relationship. *Clin Psychol Rev*, *33*(5), 686-697. https://doi.org/10.1016/j.cpr.2013.04.003

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322-338. https://doi.org/10.1037/a0014996

Ruiz, M. A., Pincus, A. L., & Schinka, J. A. (2008). Externalizing pathology and the five-factor model: a meta-analysis of personality traits associated with antisocial personality

disorder, substance use disorder, and their co-occurrence. *J Pers Disord*, *22*(4), 365-388. https://doi.org/10.1521/pedi.2008.22.4.365

Sadeghi, A., & Mozhdehi, F. K. Z. (2016). Surveying the Relationship between Personality Traits and Self-Esteem with Marital Satisfaction in Married Students of Islamic Azad University Students, City of Rasht, Guilan Province, Iran. *Psychology*, *07*(05), 655-661. https://doi.org/10.4236/psych.2016.75067

Sayehmiri, K., Kareem, K. I., Abdi, K., Dalvand, S., & Gheshlagh, R. G. (2020). The relationship between personality traits and marital satisfaction: a systematic review and meta-analysis. *BMC Psychol*, *8*(1), 15. https://doi.org/10.1186/s40359-020-0383-z

Declarations

Ethics approval and consent to participate: Does not apply

Consent for publication: Does not apply

Availability of data and materials: Does not apply

Competing interests: None

Funding: None

Authors' contributions: All material by the sole author

Acknowledgements: None