Forthcoming in *Theory and Society*.

The "Predetermined Narrative" of Methodological Rigor: A Rejoinder to Ogunbajo et al.

Callie H. Burt[1]

Abstract

This manuscript offers a rejoinder to Ogunbajo et al.'s study of "gender affirmation and adoption interest" among transgender populations and to their published reply to my critique. I argue that the study's findings are unreliable because the study's analytic strategy is not supported by its data structure. Specifically, the fitted multinomial model is applied to covariate patterns that are too sparse and imbalanced to sustain the adjusted comparisons the authors report, much less the substantive and policy claims they derive from them; key variables are operationalized in ways that conflate gender affirmation with socioeconomic advantage and misrepresent central constructs; and the article's descriptive and reporting errors further undermine confidence in its findings. At a conceptual level, I question the treatment of "adoption interest" as a public health outcome and unqualified normative good rather than as a family-formation preference whose health relevance must be argued or measured. I conclude by using this case to reflect on how epistemic norms and methodological standards are enforced—or relaxed—in contested policy domains.

Keywords: methodological critique; transgender health research; sociology of science; scientific credibility; politicization of science

[1] Callie Burt is Professor of Criminal Justice and Criminology at Georgia State University. Email: cburt@gsu.edu

Scholarly criticism is supposed to be answerable on its own terms. When a critic identifies a specific flaw in a specific analysis, the authors either show that the flaw is not there, demonstrate that it does not matter, or concede it as a flaw. What the authors are not entitled to do is leave the flaw unaddressed while recasting the critic as ignorant or ill-motivated. That substitution, of a claim about the critic for an answer to the criticism, is the central tactic in Ogunbajo et al.'s (2026b) reply to my commentary (Burt, 2026a) on their article, "Association Between Gender Affirmation and Adoption Interest Among a Large Sample of Transgender, Nonbinary, and Gender-Diverse Populations in Washington State, 2023," recently published in the *American Journal of Public Health* (AJPH).[2]

In their study, Ogunbajo et al. (2026a) modeled a three-category adoption-interest outcome with multinomial logistic regression and concluded that legal gender affirmation, health care access, economic stability, and psychological factors are associated with adoption interest among transgender individuals and carry "several important public health implications" (p.10). My comment argued that these conclusions exceeded what the data and analysis could support. My critique was methodological and conceptual: the fitted model was poorly supported by the observed data, several measures did not map onto the constructs they purported to assess, and the article drew unwarranted policy implications from a single attitudinal item: "Are you interested in adopting a child?".

The reply did not engage these points. Instead, Ogunbajo et al. characterized my comment as reflecting "either a fundamental misunderstanding of multinomial logistic regression or selective reading that prioritizes a predetermined narrative over methodological precision." The authors further warned that critiques like mine "can reinforce misinformation that unfairly undermines transgender people's right to start families." The first charge is false. My critique depends on understanding how multinomial logistic regression works and what data are required to support adjusted comparisons across outcome categories. The second charge is equally unsupported. My comment did not challenge transgender people's right to form families, access adoption, receive gender-affirming care, or obtain legal protections. My comment challenged whether this article's model, measures, and evidence could bear its conclusions. By linking my critique to misinformation and political harm, the response shifted the exchange away from methodological substance and conceptual rigor and toward motive attribution and ad hominem.

The rhetorical conversion is worth addressing. Ordinary questions about model support, measurement validity, transparency, and inference are not appropriately reframed as political

[2] This rejoinder appears here because the AJPH, which published both my original comment and Ogunbajo et al.'s reply, declined to publish a shorter rejoinder focused on the specific evidence. Beginning on May 26, 2026, upon publication of the letter and reply I wrote to the journal to raise two concerns about the reply: that it mischaracterized my methodological critique, and that it attributed improper motive, characterizing my comment as reflecting a "predetermined narrative" and situating it within a discussion of misinformation and threats to transgender people's rights. I requested an opportunity to respond in the same venue, or another remedy such as a correction or editorial note. I subsequently asked what review of these concerns the journal had conducted, who had conducted it, and whether any avenue remained for review by an editor or body not involved in the original publication decision. On July 11, 2026, the editor-in-chief replied that the journal's review was complete and that AJPH declined to publish the rejoinder; the journal did not answer the process questions. I therefore publish it here, revised and expanded.

opposition to the population under study. Moreover, scientific claims do not require less rigor because they concern a minority or vulnerable population or a salient political question.

In this rejoinder, I bring the critique and the evidence back into focus. First, I restate the methodological critique the reply did not answer: whether their model is reliable and interpretable given the sparse, highly imbalanced covariate patterns in these data. This is not a question of whether multinomial regression is permissible, nor is it a confusion of column and row percentages. Second, I show why the article's measurement choices and unclear estimand undermine the policy claims. I expand on a point that was necessarily constrained in my 400-word response, which the reply did not engage: the study examines adoption interest. It does not examine adoption applications, eligibility, readiness, encountered barriers, discrimination in adoption systems, unmet need, or movement from interest to action. In sum, the study cannot support the public health and policy claims the article advanced.

The model is unreliable

My central critique was model reliability and robustness, namely, whether the data contain enough within-category variation across covariate patterns to support the adjusted multinomial estimates. They do not. A fully adjusted multinomial model estimates coefficients jointly from the observed covariate structure. When several predictors are sparse, highly imbalanced, or nearly deterministic within an outcome category, as they are here, the model has little basis for separating their adjusted associations, and the resulting estimates are unstable and difficult to interpret even when the analysis converges (Agresti, 2013; Long, 1997).

The sparseness of the data structure is displayed in their Table 1. The "interested-in-adoption" category contains 440 of 788 respondents, and within it, 425 are transgender women against 11 transgender men and 4 nonbinary individuals; 426 respondents were coded assigned male at birth against 14 coded as assigned female; 436 respondents are employed against 4; 433 respondents are above the poverty line against 7. Among those who wanted the relevant care, 12 of 439 lacked hormone therapy and 8 of 435 lacked surgery.[3] Meanwhile, the final model estimates more than two dozen (~27) slope coefficients for each of the two multinomial contrasts (interested versus not interested, and might-be interested versus not interested). Fitting that many parameters across three outcome categories, when the informative variation on several predictors consists of four, seven, or eleven people, necessarily forces many covariate-pattern cells to hold one or two observations and some to hold none. Crucially, this affects the model as a whole, not just estimates for sparse cells. As noted, multinomial models estimate all coefficients jointly from a single likelihood, and near-constant predictors produce quasi-separation that destabilizes the whole estimation (Albert & Anderson, 1984). Therefore, no estimate can reasonably be read as reliable. This is a property of the design applied to this sample, not a matter of interpretation (Agresti, 2013; Long, 1997). The reply's emphasis on variation in marginal distributions—for example, across race, geography, and

[3] Ogunbajo et al.'s reply notes that peer reviewers raised plausible reasons to expect adoption considerations to differ for transgender men (differential reproductive anatomy and the potential to give birth themselves), which the authors addressed in their discussion. A theoretical rationale for subgroup differences, however, does not supply the data needed to estimate those differences reliably, and it does not remedy the sparsity that leaves the transgender men and nonbinary strata effectively uninformative in the fitted multinomial model.

education—does not address this concern, because multinomial estimates depend on the sparse *joint* distribution of covariates and outcomes rather than on heterogeneity in single predictors taken one at a time.

The confidence intervals are signatures of the problem. The adjusted odds ratio for having discussed fertility options with a provider in the interested-versus-not-interested contrast is 19.2, with a 95% interval of 2.38 to 154.51, an interval whose upper bound is roughly *sixty-five times* its lower bound. Private versus public insurance in the same contrast returns 11.01 (2.89 to 42.02). The odds of being interested in adoption (versus not) for transgender men relative to transgender women is 2.53 (0.47 to 13.55), spanning a halving to a thirteen-fold increase in odds. Intervals of this width signal unreliability.

Ogunbajo et al. note in their article and again in the reply that they applied "nonparametric bootstrapping with 1000 iterations … to address potential model instability from the modest sample size." But bootstrapping cannot remedy an instability that is rooted in the data rather than in the estimator (Kosmidis & Lunardon, 2024). Resampling draws repeatedly *from the same sparse observations*; it therefore cannot supply information the sample does not contain. Where a covariate pattern holds sparse observations or none, no number of iterations will manufacture the variation needed to estimate its coefficient. The extremely wide confidence intervals are the result. Effectively, this bootstrapping procedure dresses up the estimates without adding the information they lack.

As to model fit, in the article, the authors cite a pseudo-R^2 of 0.28 as evidence of "adequate model performance," but a pseudo-R^2 is uninformative about the precision of the individual coefficients on which every substantive and policy claim depends. Global fit and coefficient stability are different quantities, and the former is what is lacking here.[4]

The column-percentage rebuttal, reversed

The reply's one attempt at a methodological rebuttal is a claim that I misread column percentages as row percentages. I did not. In their response, Ogunbajo et al. note that while 96.6% of those interested were transgender women, transgender women are only 82.2% of the total sample, and they pivot to row percentages (65.6% of transgender women interested, versus 14.7% of transgender men and 6.0% of nonbinary respondents) as though these answered my composition point.

[4] Pseudo-R^2 (McFadden's, which Stata reports and is almost certainly the 0.28 here) is a function of the fitted log-likelihood versus the null-model log-likelihood. It is a global summary of how much the predictors collectively improve fit over an intercept-only model. Coefficient stability is a per-parameter property, driven by how much informative variation stands behind each estimate. A model can have a respectable pseudo-R^2 while individual coefficients are wildly unstable, which is this model. A few strong, well-populated predictors (or, more perversely, the near-separation itself) can drive the likelihood up and inflate the pseudo-R^2, while sparse cells leave specific coefficients with enormous standard errors. Indeed, quasi-complete separation can push a pseudo-R^2 toward its ceiling because the model fits the sample almost perfectly, at the same time as it sends the affected coefficients and their intervals toward infinity, thus producing high global fit and unreliable local estimates.

They do not answer my point, and the rebuttal fails on inspection. My claim concerned the composition of the *interested-in-adoption category*, which is a column percentage, and the correct column percentage is 96.6%, as I noted. That transgender women are a large share of the whole sample is a different quantity and does not address how homogeneous the "interested-in-adoption" cell is. The pivot to row percentages, moreover, simply reintroduces the sparse subgroups (11 transgender men, 4 nonbinary respondents) which drive the instability described above.

The irony is that the authors committed the error they attribute to me. On page 3 of their article, Ogunbajo et al. state that "82.2% [of those] interested in adoption were transgender women." The interested column shows 425 of 440, which is 96.6%. As noted, the 82.2% figure is the whole-sample share of transgender women (648/788) imported into a statement about the interested group; thus, Ogunbajo et al. commit the column-versus-row confusion they accuse me of. Thus, the only passage in which their reply engages the statistics exposes the error they projected onto me.

Conditioning, construct validity, and the estimand

My second objection concerned how the affirmation variables were constructed. Medical and legal affirmation were assessed only among respondents "whose gender goals included" the relevant interventions. The reply defends this as standard participant-centered practice, noting that it would be inappropriate to ask whether someone received hormone therapy they never wanted. I do not disagree, and I did not dispute that.

My critique was that conditioning the affirmation measures on a self-selected subgroup changes what is estimated and, in this instance, loads the measures with socioeconomic content. Access to hormones, surgery, name changes, and gender-marker updates is structured by insurance, income, time, proximity to providers, and legal assistance. A variable recording receipt of such care among those who sought it is, in substantial part, a proxy for the resources required to obtain it. Entered alongside insurance, employment, poverty, and education, themselves imbalanced and collinear, this leaves the model unable to separate "affirmation" from "advantage," and the proposed mechanism becomes empirically indistinct from the simpler account that materially advantaged respondents both obtain care and express interest in a costly family-formation pathway. Their reply answers the survey-design question and neglects the inferential one.

There is also a major reporting problem. The article reports a total sample of 788 and a final complete-case model of 756.[5] However, because the authors only assessed medical and legal affirmation among respondents whose "gender goals" included them, these variables are defined on far smaller analytic subsets of the sample. The number of respondents who wanted surgery appears to be 602 (reported in their Table 1 as 609, see Appendix). A complete-case model that includes receipt of surgery only "if wanted" thus cannot exceed 602 respondents, and would be smaller after other missingness. Their reported model N of 756 is therefore not reconcilable with

---

[5] The authors write: "We employed a comprehensive case analysis to develop models." A comprehensive case analysis is not a term with an established meaning in the missing-data literature. I assume given the context that this was a malapropism and that complete-case analysis (i.e., list-wise deletion) was used. If something else is intended, the paper does not say what.

the stated coding. Notably, this is not a trivial reporting matter. Gender affirmation is the central exposure, and legal affirmation (the updated-name variable), supplies its headline association (adjusted odds ratio 4.7). Either the medical-affirmation variables are not in the model as described, or non-wanters were folded in through an undisclosed coding decision that again shifts the estimand. Either way, the reader cannot tell what population the affirmation coefficients describe.

Two additional operationalization limitations compound the issues, and the reply addresses neither. "Internalized transphobia," a construct with validated multi-item instruments, is operationalized through a single item asking whether respondents are glad to be trans. A single positively worded item does not capture internalized stigma. By analogy, you can accept yourself as a gay or left-handed person, without "being glad" you are gay or left handed. Additionally, the model nonsensically includes both overall family support and the proportion of supportive family members. These are overlapping variables whose joint inclusion asks the model to estimate the effect of the proportion of family support net of the level of family support, a quantity that is conceptually incoherent and that invites the instability already described. Silence on both is notable, because these are the objections least amenable to reframing: they concern what was measured, and no appeal to expertise or number of authors answers them.

Association is not the issue

The reply's seemingly most effective response is to cast my critique as the misapplication of causal-inference standards to avowedly descriptive research, noting that the authors stated the cross-sectional design "prevents causal inference" and used associational language throughout.

This mischaracterizes my objections. Sparse-cell instability, uninterpretably wide intervals, single-item constructs, and collinear predictors are problems of estimation and measurement, and they are equally fatal to descriptive claims. An adjusted odds ratio of 19.2 with an interval from 2.38 to 154.51 carries little usable information about the association it purports to describe, and relabeling the coefficient "associational" rather than "causal" neither narrows the interval nor populates the empty cells. The association-versus-causation distinction is real and important; however, it is fully orthogonal to the concerns I raised.

The same applies to the interpretive overreach in the policy claims. A single attitudinal item, "Are you interested in adopting a child?," measures interest and nothing more. The paper nonetheless concludes that its findings "underscore the need for coordinated efforts across health care, legal, and social service sectors to create more inclusive and supportive pathways to family formation," a claim that presupposes unmet need and systemic exclusion, neither of which a measure of interest can establish. The reply answers that "descriptive epidemiology has long informed public health practice," which is true and general. Descriptive findings can motivate hypotheses about need; they cannot, on their own, demonstrate the specific barriers and exclusions their policy recommendations concern. My objection was not that description cannot inform practice, of course, but that this description does not support these particular claims.

When the descriptive results fail basic checks

The concerns above are illustrative, not exhaustive, and they sit atop other errors that raise concerns. The descriptive results contain at least eight demonstrable inconsistencies, catalogued in the Appendix. These include counts that cannot sum to their column totals, a percentage off by an order of magnitude due to a typo, a subgroup denominator reported two different ways, and text figures drawn from the wrong column or the wrong outcome group. Individually, any one might be excused as a production error. Collectively, they show that the tables and text were not reconciled against one another or against the underlying data before publication. When the visible descriptive presentations are this unreliable, there is little basis for assuming that the far less visible modeling decisions were executed and reported with greater care, and the reader is left unable to verify that the quantities feeding the multinomial model are the ones the authors claim.

Similar issues can be found in the interpretation. In their conclusion, the authors list policy implications that allegedly follow from their study findings. One is: "mental health services specifically addressing internalized transphobia and parenting concerns should be integrated into adoption services and made available during the early stages of family planning" (p.10). Although a stretch, a finding that internalized transphobia is negatively associated with adoption desire is at least somewhat consistent with this idea. That is *not what the results show*, however. Instead, the internalized transphobia measure is not significantly associated with interest versus no interest. However, it is significantly associated with maybe interested versus no interest, in the opposite direction. Individuals who score higher on their internalized transphobia measure are more likely to be maybe interested than not interested in adoption. If you take it at face value, internalized transphobia is associated with more openness to adoption than the reference group, not less. The clinical recommendation implicitly assumes internalized transphobia is a barrier to be treated so people can move toward family formation, but their own estimate points the other way (or does not support the barrier story at all).

What is the health explicandum?

The concerns catalogued to this point concern how the study was carried out: how the model was fit, how the variables were built, whether the tables reconcile. Another issue worth discussing concerns the explicandum: why is adoption interest a public health outcome? The dependent variable is a single item asking whether the respondent is interested in adopting a child. That is a stated interest in one route to family formation. It is not a health state, a health behavior, a clinical outcome, or the use of a health service, and the paper offers no account of how it proxies any of these. Adoption interest is nowhere shown to index health, to be produced by health, or to produce it. The variable enters the analysis as though its standing as a health outcome were self-evident, and it is not. It is clear that the paper takes a normative stance toward higher adoption interest as good, but on what scientific or theoretical basis? The model implicitly treats higher adoption interest—versus interest in biological parenthood, no children, or already having children—as beneficial, without explaining why this particular preference should be privileged as a public health goal.

In a thorough model of family formation, adoption interest would be situated among other family-formation states and preferences—having children already, wanting biological children, not wanting children—rather than treated as a free-standing outcome, and the paper does not locate the item within such a framework. However, even that framing still does not make adoption interest, per se, a *health* issue.

In my reading, the health framing is accepted not because of the health-relevance of the topic but the population. Adoption interest becomes a "public health" outcome because the respondents are transgender, as though studying a minority or vulnerable group was enough to convert any attitude about them into a health matter with policy implications. It is not. Vulnerability does not lower the bar for what counts as a health outcome or supply a health mechanism. The same item, fielded without the vulnerable-population frame (e.g., adoption interest among non-transgender populations) would almost surely be recognized as a preference rather than a health outcome, and nothing about the respondents' being transgender changes what the item measures. Invoking a vulnerable population is not a substitute for establishing health relevance, and this study's health-framing relies on that substitution.

The outcome is also weak on the narrower ground of what "interested" means. Answered yes, no, or maybe, with no cost, timeline, or alternative specified, the item records a disposition in the abstract rather than a preference under real constraints. The article could have asked whether respondents wanted or intended to adopt, and instead asked whether they were "interested," a word that commits to neither curiosity, desire, nor intention and that the paper never defines. Of course, that would still not make it a health outcome.

None of this is pedantic boundary questioning, because the article's policy apparatus rests on it. The conclusions call for provider training, health-system initiatives, integrated mental-health support, and economic assistance, all interventions in the health and social-service sectors, and those follow only if adoption interest *is* a health matter with a health mechanism to act upon. Absent an argument for why *greater adoption desire itself* advances health or equity, the recommendation to promote it amounts to endorsing one family-formation route over others on unstated normative grounds.

The routes that might supply a health warrant are possibly available in principle but are left unmade. A wellbeing framing would require showing that adoption interest tracks wellbeing, which the study does not test. A reproductive-justice framing treats family-building access as a domain of health equity, but that is an argument about barriers and access, and the study measures neither. The authors recognize: "Our focus on individual-level characteristics may not fully capture systemic barriers in adoption systems or variations in state-level policies that affect trans individuals' adoption experiences" (p.10). Not fully capture is doing illegitimate work here. The study did not partially capture systemic barriers to adoption; it captured none. There is no barrier variable in the dataset. No measure of agency policy, no state-level variation, no discrimination experienced in adoption processes, no application, no attempt, no denial. This invalidates the paper's conclusions, which are about barriers—inclusive pathways, systemic exclusion, coordinated efforts across health care, legal, and social service sectors. These are all claims about things the study did not assess.

The defense the reply leans on, descriptive epidemiology of an understudied population, still requires that the quantity described be health-relevant; describing a population's interest in adoption is descriptive social science unless the health link is argued, and here it is not. The result is that the study's "public health implications" are asserted rather than demonstrated with a normative overlay. The outcome is of unestablished health relevance, mismatched to the family-formation conclusions it is marshaled to support, and vague in its own terms, and the policy conclusions borrow a health warrant the analysis never earns.

The rhetoric that replaced the answers

Set the evasions side by side and a pattern emerges—replacement of substantive engagement with rhetoric. Of my substantive objections, one is answered in a way that supports the objection, one is answered by addressing a question I did not ask, and several are not answered at all. In place of the missing answers, the reply supplies three rhetorical dodges, and it is worth calling them out.

The first is the attribution of motive. The reply offers the reader a choice of two explanations, and both are about me rather than my argument: that I have "a fundamental misunderstanding" of the model or have engaged in a "selective reading that prioritizes a predetermined narrative," and the second imputes bad faith without evidence. I raised checkable technical points. The appropriate response is to check and address them.

The second is the appeal to authority. The reply emphasizes that its authors are "an interdisciplinary team of researchers with decades of experience," and frames disclosure of expertise as strengthening the critique. A coefficient estimated from four observations is not stabilized by the seniority of the analyst who reports it or decades of experience (which I also possess). Expertise is not the currency here, and it is not the province of one side. The evidence is available to anyone who checks.

The third is the invocation of political stakes, the warning that critiques like mine "can reinforce misinformation that unfairly undermines transgender people's right to start families." This argument attempts to shield this work from critique by implying that identifying methodological weakness in a study of transgender people is a harmful political act, and is, therefore, something a responsible scholar should decline. I reject that premise. Fragile findings are not protective. An odds ratio of 19.2 with an interval to 154 is a result that collapses under adversarial scrutiny, and a literature built on such results is more, not less, vulnerable. Scholarship has identified a troublesome pattern of lax scrutiny toward more ideologically congenial findings (e.g., Clark & Winegard, 2020; Honeycutt & Jussim, 2020; Savolainen, 2024; Tetlock, 1994), and the idea that lax scrutiny to ideologically congenial scientific claims benefits certain protected groups is as misguided as it is myopic (Burt, 2026b).

Finally, and most minor, there is the matter of my name. Across two published pages, the authors refer to me only as "the commentator," never as Burt, in a signed exchange in which my letter appears under my name and theirs under theirs. The effect is to depersonalize the critic while personalizing the charge against her: the individual is reduced to a role, and the role is assigned a "predetermined narrative." I note it not because a name is owed to me as a courtesy, although it is, but because the omission is of a piece with the rest of the reply. It is surely easier to impute an agenda to "the commentator" than to a named colleague, and easier still when the argument being answered has been left on the page unaddressed.

Conclusion

I want to be clear about what I am and am not claiming. I am not claiming that adoption interest among transgender people cannot be studied, that descriptive research is illegitimate, or that the authors acted in bad faith. I am claiming that this model, applied to this sample, produces estimates too unstable to interpret; that the affirmation measures conflate affirmation with material advantage in ways the paper does not resolve; that the reported analytic sample cannot be reconciled with the stated coding; that two measures do not capture their stated constructs; and that a single item measuring interest cannot bear the policy conclusions drawn from it. These claims are answerable, and the reply did not answer them.

My broader point concerns how a field polices its own quality. Methodological standards are not a predetermined narrative. They are the conditions under which empirical claims earn confidence and through which (robust) knowledge accumulates. When methodological criticism in a contested domain is met with attributions of ignorance and bad faith, appeals to seniority, and warnings about political harm, the immediate effect is to protect one paper from one critic. The lasting effect is to teach everyone watching that critique will be recoded as hostility, and that silence is the safe course (Burt, 2022). A literature that cannot be criticized cannot be trusted, and a population whose defenders rest their case on unreliable findings has been poorly served.

References

Agresti, A. (2013). *Categorical data analysis* (3rd ed.). Hoboken, NJ: Wiley.

Albert, A., & Anderson, J. A. (1984). On the existence of maximum likelihood estimates in logistic regression models. *Biometrika, 71*(1), 1–10.

Burt, C. H. (2022). Discounting females, denying sex, and disregarding dangers from self-ID—A reply and a defense of open debate. *Journal of Controversial Ideas, 2*(1), 1–20.

Burt, C. H. (2026a). On adoption interest among transgender individuals: Conceptual ambiguity, methodological weakness, and interpretive overreach. *American Journal of Public Health, 116*(7), 887–889.

Burt, C. H. (2026b). Ideological homogeneity and its epistemic discontents: A case study of research on transgender issues. *Theory and Society, 55*(3), Article 48.

Clark, C. J., & Winegard, B. M. (2020). Tribalism in war and peace: The nature and evolution of ideological epistemology and its significance for modern social science. *Psychological Inquiry, 31*(1), 1–22.

Honeycutt, N., & Jussim, L. (2020). A model of political bias in social science research. *Psychological Inquiry, 31*(1), 73–85.

Kosmidis, I., & Lunardon, N. (2024). Empirical bias-reducing adjustments to estimating functions. *Journal of the Royal Statistical Society Series B: Statistical Methodology, 86*(1), 62–89.

Long, J. S. (1997). *Regression models for categorical and limited dependent variables* (Vol. 7). Thousand Oaks, CA: Sage.

Ogunbajo, A., DeAtley, T., Breslow, A., Gamarel, K., Operario, D., Kershaw, T., Emery Tavernier, R. L., & Restar, A. J. (2026a). Association between gender affirmation and adoption interest among a large sample of transgender, nonbinary, and gender-diverse populations in Washington State, 2023. *American Journal of Public Health, 116*(2), 229–239.

Ogunbajo, A., DeAtley, T., Breslow, A., Gamarel, K., Operario, D., Kershaw, T., & Restar, A. J. (2026b). Ogunbajo et al. respond. *American Journal of Public Health, 116*(7), 891–892.

Savolainen, J. (2024). Unequal treatment under the flaw: Race, crime & retractions. *Current Psychology, 43*(17), 16002–16014.

Tetlock, P. E. (1994). Political psychology or politicized psychology: Is the road to scientific hell paved with good moral intentions? *Political Psychology, 15*(3), 509–529.

Appendix: Errors in the Descriptive Results

The descriptive results contain a series of errors. In Table 1, multiple cells fail basic internal checks, including counts that cannot sum to their column totals, apparent typos, and a subgroup denominator reported two different ways. The text compounds the problem, citing figures drawn from the wrong column or the wrong outcome group on at least two occasions. They are catalogued below.

- **Table 1, education, "≤ high school," interested column.** The cell prints 365, but the interested column must total 440, and 365 + 161 + 14 = 540, which is impossible. The reported 60.2% identifies the intended value, since 265/440 = 60.2% and 265 + 161 + 14 = 440. The cell should read 265; the leading digit was apparently mistyped as a 3 instead of a 2.

- **Table 1, sex assigned at birth, "assigned male at birth," not-interested column.** The cell posts 105, which forces the row to total 662 against the stated 659 and the not-interested column to total 164 against the required 161. The reported 63.4% fixes the value, since 102/161 = 63.4%. The cell should read 102 (also corroborated by the poverty-level "No" not-interested cell, which is 102 and also reported as 63.4% on the same denominator).

- **Table 1, fertility discussion, "No, and do not want to," not-interested column.** The count of 130 is correct, but the percentage prints as 9.3% when 130/144 = 90.3% on the variable's not-interested denominator of 144. A zero was dropped, turning 90.3% into 9.3%.

- **Table 1, gender identity, "nonbinary," not-interested column.** The count of 33 is correct, but the percentage prints as 19.3% when 33/161 = 20.5%.

- **Table 1, gender-affirming surgery "if wanted," denominator.** The row header states n = 609, but the cells sum to 602 (88 + 514, and group cells 85 + 435 + 82), and every percentage in the row is computed on 602. The denominator is reported two ways, off by seven. This does not undermine the model-sample objection; if anything, the ceiling for a complete-case model including this variable drops to ≤ 602.

- **Text–table contradiction, page 3, gender identity of the interested group.** The text states that 82.2% of those interested in adoption were transgender women, but the interested column shows 425 of 440, or 96.6%. The 82.2% figure is the overall-sample share of transgender women (648/788) pulled from the wrong column. The interested group is 96.6% transgender women, a more extreme concentration than the text reports (and consistent with my figure).
- **Text–table contradiction, page 3, adoption interest among those coded as 'assigned female at birth'.** The text states that 44.2% of female individuals were interested in adoption, but the interested female cell is 14 of 122, or 11.5%. The 44.2% corresponds to the not-interested rate (54/122), so the figure is drawn from the wrong outcome group. The correct interested rate for those female respondents is 11.5%, the counterpart to the correctly stated 64.6% for male individuals.
- **Presentation issue, Table 2, age row.** The model reports age as "0.91 ± 0.82" and "0.92 ± 0.85" under a header reading "AOR (95% CI) or Mean ± SD." A continuous covariate in a multinomial logistic model should be reported as an adjusted odds ratio with a confidence interval, not a mean and standard deviation, which is uninterpretable as a model estimate.

Again, individually, any one of these might be excused as a transcription error in production. Collectively, they reveal that the tables and text were not reconciled against one another or against the underlying data before publication. When the descriptive results contain this many demonstrable inconsistencies, there is little basis for assuming the far less visible modeling decisions were executed and reported with greater care, and the reader is left unable to verify that the quantities feeding the multinomial model are the ones the authors claim.

Conflict of Interest Statement: The author has no conflicts of interest to report.

Funding: The author received no financial support for this work.