A Review of "Specification Curve Analysis Shows That Social Media Use Is Linked to Poor Mental Health, Especially Among Girls" by Twenge et al. (2022)

Reviewer 2

January 14, 2026

v1

isitcredible.com

Disclaimer

This report was generated by Reviewer 2, an automated system that uses large language models to assess academic texts. It has been read and approved by a human editor on behalf of The Catalogue of Errors Ltd. The report's goal is to facilitate the discovery of knowledge by identifying errors in the existing literature. Comments can be made here. Any errors will be corrected in future revisions.

> I am wiser than this person; for it is likely that neither of us knows anything fine and good, but he thinks he knows something when he does not know it, whereas I, just as I do not know, do not think I know, either. I seem, then, to be wiser than him in this small way, at least: that what I do not know, I do not think I know, either.
>
> Plato, *The Apology of Socrates*, 21d

> To err is human. All human knowledge is fallible and therefore uncertain. It follows that we must distinguish sharply between truth and certainty. That to err is human means not only that we must constantly struggle against error, but also that, even when we have taken the greatest care, we cannot be completely certain that we have not made a mistake.
>
> Karl Popper, 'Knowledge and the Shaping of Reality'

Overview

Citation: Twenge, J. M., Haidt, J., Lozano, J., & Cummins, K. M. (2022). Specification Curve Analysis Shows That Social Media Use Is Linked to Poor Mental Health, Especially Among Girls. *Acta Psychologica*, Vol. 224, 103512.

URL: https://doi.org/10.1016/j.actpsy.2022.103512

Abstract Summary: The current paper re-ran a Specification Curve Analysis (SCA) on adolescent technology use and mental health data using revised constraints, finding that social media use is consistently and substantially associated with poor mental health, particularly among girls. These associations were much larger than previously reported and stronger than links between mental health and several high-risk behaviors like binge drinking and hard drug use.

Key Methodology: Specification Curve Analysis (SCA) applied to three large-scale community datasets (Millennium Cohort Study, Monitoring the Future, Youth Risk Behavior Surveillance System).

Research Question: How does the median Beta change when analyses zero in on girls and social media, exclude potential mediators as control variables, and treat all mental health scales equally, compared to the original SCA findings by Orben and Przybylski (2019)?

Summary

Is It Credible?

The article presents a forceful challenge to the prevailing consensus established by Orben and Przybylski (O&P) regarding the minimal association between adolescent technology use and mental health. By re-analyzing three large datasets using Specification Curve Analysis (SCA), the authors argue that the link is "consistent and substantial," particularly for girls using social media, with effect sizes comparable to those of binge drinking or sexual assault. While the authors successfully demonstrate that analytical choices can dramatically alter results, their conclusion that the association is "substantial" relies entirely on a specific, disputed causal assumption. The credibility of the findings is therefore not established by the data itself, but by the reader's willingness to accept the authors' theoretical premises regarding control variables.

The central issue lies in the treatment of "potential mediators." The headline finding—a median beta of -0.20 for girls in the Millennium Cohort Study (MCS)—materializes only when a specific set of control variables used by O&P (such as closeness to parents and attitudes toward school) are excluded. The authors argue these are mediators (downstream effects of social media) and removing them prevents "overadjustment bias." However, if these variables are confounders (e.g., family distress causes both social media escape and poor mental health), excluding them introduces severe Omitted Variable Bias. The article reveals that including these controls reduces the effect size from -0.20 to -0.01. Consequently, the "substantial" link is not a robust feature of the data; it is an artifact of attributing nearly all shared variance between social media and psychosocial distress to the causal influence of social media. This is a theoretical gamble that cross-sectional data cannot validate.

Furthermore, the robustness of the findings is not fully supported by the data, given the sensitivity to the exclusion of controls. The authors present their results as the product of four "theoretically defensible" constraints. However, the data indicates that the exclusion of mediators is the dominant driver of the effect size, suggesting the effect is not robust across all four constraints equally. Additionally, the decision to "treat scales equally" in

the MCS dataset effectively upweights self-report measures (which share method variance with the exposure variable) while downweighting the parent-report Strengths and Difficulties Questionnaire (SDQ), which showed negligible effects. By shifting the analysis toward self-reports and removing psychosocial controls, the methodology maximizes the potential for inflating associations through Common Method Variance and unmeasured confounding.

Finally, there are inconsistencies in how the authors apply their own standards. While they criticize O&P for lumping "screen time" activities together, they commit a similar error in the YRBSS dataset by using "electronic device use" (which includes gaming) as a proxy for social media, yet labeling it "Social media" in their summary figures. Moreover, the authors infer a causal role for social media in the longitudinal rise of mental health issues since 2012 based on these cross-sectional correlations. Given the extreme sensitivity of the effect size to the inclusion of controls, the claim that these associations are "practically important" enough to explain population-level trends is an overreach. The article ultimately proves that SCA is highly sensitive to researcher degrees of freedom, but it does not definitively establish that the "true" association is the substantial one they prefer.

The Bottom Line

The study effectively demonstrates that the link between social media and mental health can appear substantial if one assumes that social media causes all co-occurring psychosocial problems (like family distance or school dislike). However, the data shows that controlling for these factors reduces the association to near zero. Because the analysis relies on cross-sectional data that cannot distinguish between a confounder and a mediator, the claimed "substantial" effect is a reflection of the authors' theoretical choices rather than indisputable evidence from the datasets.

Specific Issues

Omitted variable bias and mediator exclusion: The study's primary conclusion hinges on the decision to exclude control variables such as "closeness to parents" and "school grades," labeling them as mediators rather than confounders. The article shows that including these variables reduces the median beta for girls in the MCS dataset from -0.20 to -0.01 (p. 4). This massive reduction implies that these variables account for approximately 95 percent of the covariance. By excluding them, the authors assume the causal direction flows entirely from social media to these psychosocial factors. If this assumption is incorrect, the reported effect sizes suffer from critical Omitted Variable Bias.

Narrative obscuring of constraint impact: The authors frame their analysis as a package of four equal methodological improvements. However, the results demonstrate that the exclusion of mediators is the single necessary condition for the large effect sizes. Even when applying the other constraints—separating sexes, isolating social media, and treating scales equally—the inclusion of the disputed controls collapses the effect to -0.01 (p. 4). The narrative does not sufficiently emphasize the fact that the result is not robust to standard psychosocial controls.

Scale weighting: In the MCS analysis, the constraint to "treat scales equally" explicitly reduces the influence of the parent-report SDQ (which showed lower betas) and increases the weight of self-report measures (p. 3). This effectively triples the influence of measures that share a rater with the exposure variable, thereby amplifying Common Method Variance. The "substantial" finding is partly driven by prioritizing self-reports over parent-reports.

Mislabeling of variables in YRBSS: The authors criticize previous research for lumping screen activities together but use "electronic device use" (which includes gaming) as a proxy for social media in the YRBSS dataset (p. 6). Despite acknowledging this limitation, the summary Figure 6 explicitly labels this data point under "Social media," visually misrepresenting the specificity of the finding and violating the authors' own constraint regarding activity separation (p. 9).

Causal overclaiming from cross-sectional data: The authors suggest it is "plausible" that

digital media use is responsible for the longitudinal rise in adolescent depression since 2012 (p. 10). This infers a longitudinal cause from cross-sectional correlations. Furthermore, benchmarking these correlational effect sizes against causal interventions, such as lead removal ($r = -0.09$, p. 9), implies a level of causal certainty and utility that the cross-sectional analysis does not support.

Methodological inconsistency ("Dueling SCAs"): The authors criticize O&P for using researcher degrees of freedom to minimize effects, yet their own analysis utilizes degrees of freedom (selecting girls, excluding mediators, weighting self-reports) that result in maximized effect sizes. Rather than a neutral SCA that aggregates all defensible specifications to show the full range of uncertainty, the article presents a selective slice of the specification curve that supports the authors' hypothesis.

Data contradicting the "social media" hypothesis: The authors argue that social media is distinct from other screen time in its harm. However, in the MCS dataset, "Internet use" shows a slightly stronger negative association ($\beta = -0.21$) than social media ($\beta = -0.20$) for girls (Table 1a, p. 4). This suggests the specific mechanism of "social media" may not be distinct from general internet usage, complicating the article's third contribution.

Unexplained insensitivity to controls in YRBSS: In the YRBSS analysis, the beta remains exactly -0.11 regardless of whether no controls, race controls, or race and grade controls are used (Table 3a, p. 8). While the authors attribute this to the lack of mediator variables in the dataset (p. 6), the complete lack of variation (0.00 change) across different demographic specifications is unusual.

Measurement limitations in MtF: The Monitoring the Future (MtF) analysis relies on a single-item measure of happiness with a 3-point scale (p. 6). While the authors acknowledge this limits the SCA to a single specification, relying on such a low-resolution outcome variable weakens the claim of "practical significance" derived from this dataset.

Selective reporting of associations: The narrative focuses heavily on the "consistent and substantial" negative associations found among girls using the authors' preferred constraints. Null or positive associations, such as those found for boys or when using O&P's controls (e.g., +0.04 for boys in Table 1b, p. 4), are reported in tables but receive

less emphasis in the text than the negative associations, thereby supporting the specific "especially among girls" argument.

Procedural and presentation issues: There are minor issues regarding variable definitions and presentation. The MtF analysis introduces a "hours per week" variable without descriptive statistics, and the Figure 6 legend incorrectly refers to "hours per day" for the MtF data (p. 9). Additionally, the comparison with O&P's YRBSS results is complicated by the use of differing time periods (p. 6).

Future Research

Longitudinal cross-lagged panel models: To address the critical issue of directionality between social media use and "mediator" variables (e.g., school attitudes, closeness to parents), future research must employ longitudinal designs. Cross-lagged panel models could help determine whether social media use precedes family distancing or vice versa, thereby validating or refuting the decision to exclude these controls.

Multiverse analysis of causal assumptions: Instead of "dueling SCAs" where researchers select their preferred constraints, future analysis should perform a multiverse analysis that explicitly models the results conditional on different causal assumptions. This would quantify exactly how much the estimated effect size depends on the classification of variables as confounders versus mediators, rather than obscuring this dependency.

Objective digital trace data: To resolve issues related to self-report measurement and the lumping of activities (e.g., "electronic device use"), research should utilize objective logging of digital activity. This would allow for a precise distinction between social media, gaming, and general internet use, and eliminate the Common Method Variance introduced by self-reporting both exposure and mental health outcomes.