A Report on "Specification Curve Analysis Shows that Social Media Use is Linked to Poor Mental Health, Especially Among Girls" by Twenge et al. (2022)

Reviewer 2

February 05, 2026

v1

isitcredible.com

Disclaimer

This report was generated by large language models, overseen by a human editor. It represents the honest opinion of The Catalogue of Errors Ltd, but its accuracy should be verified by a qualified expert. Comments can be made here. Any errors in the report 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*, 224, 103512.

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

Abstract Summary: This paper re-runs a Specification Curve Analysis (SCA) on three large-scale adolescent datasets, applying revised constraints (separating media types, separating sexes, excluding mediators, and treating scales equally) to re-examine the link between technology use and mental health. The revised analysis finds a consistent and substantial association between social media use and poor mental health, particularly among girls, contradicting previous findings that suggested the association was tiny.

Key Methodology: Specification Curve Analysis (SCA) re-run on three large-scale community datasets (Millennium Cohort Study, Monitoring the Future, Youth Risk Behavior Surveillance System) using revised analytic constraints.

Research Question: Does the association between adolescent technology use and mental health remain tiny when Specification Curve Analysis (SCA) is re-run using alternative, theoretically defensible analytic constraints?

Summary

Is It Credible?

This article presents a forceful rejoinder to the prevailing narrative that the link between digital media use and adolescent mental health is negligible. By re-analyzing three large datasets previously examined by Orben and Przybylski, Twenge et al. argue that the earlier findings of "tiny" associations (median betas between -0.01 and -0.04) were artifacts of analytical choices that diluted the signal. The authors contend that when the analysis focuses specifically on girls, isolates social media from general screen time, treats mental health scales equally, and—crucially—removes control variables they identify as mediators, the association becomes "consistent and substantial" (p. 1). They report median betas around -0.20 for girls, an effect size they demonstrate is comparable to the associations between mental health and serious risk behaviors like binge drinking and sexual assault (p. 5).

The credibility of this "substantial" finding, however, rests precariously on a single, highly contestable methodological decision: the exclusion of certain control variables. The authors classify variables such as "negative attitudes toward school," "closeness to parents," and "school grades" as potential mediators—downstream consequences of social media use—and therefore exclude them to avoid overadjustment bias (p. 2). This decision is the primary lever moving the results. In the Millennium Cohort Study analysis for girls, the median beta is -0.01 when these controls are included, but jumps to -0.20 when they are removed (p. 4). The article offers no empirical evidence or formal causal model to substantiate the claim that social media use causes school unhappiness or family distance, rather than these factors acting as common causes (confounders) that drive both social media use and poor mental health. If these variables are actually confounders—for instance, if a difficult home environment leads a teen to seek refuge in social media *and* suffer mentally—

then the authors' preferred model is misspecified, and the "substantial" association is largely spurious.

Furthermore, while the article effectively critiques Orben and Przybylski for lumping distinct activities into "screen time," it faces its own measurement validity issues. In the analysis of the Youth Risk Behavior Surveillance System (YRBSS), the authors claim to find support for the social media hypothesis, yet the dataset lacks a specific social media question. Instead, they rely on a measure of "electronic device use" that explicitly includes "computers and gaming consoles" (p. 6). Given the authors' own finding in other datasets that gaming has weaker associations with mental health than social media, using this broad proxy to support a specific claim about social media is methodologically inconsistent (p. 4). It weakens the assertion that the findings are robust across all three datasets.

The framing of the results also warrants scrutiny. The authors situate their findings within the "mystery" of the sudden rise in adolescent depression around 2012, implying a causal role for the concurrent rise of social media (p. 1). While they use correlational language in the results, the narrative structure strongly suggests causation. However, the cross-sectional nature of the data means it cannot distinguish between social media harming mental health, poor mental health driving social media use, or unmeasured third variables affecting both. The article acknowledges that "social media should remain on the list of possible explanations," but the strength of the causal inference implied by the "mystery" narrative outstrips what these specific data can support (p. 10).

Despite these limitations, the article makes a significant contribution by demonstrating the sensitivity of Specification Curve Analysis (SCA) to researcher decisions. It successfully dismantles the idea that the Orben and Przybylski analysis provided a definitive "final word" (p. 2). By showing that theoretically defensible alternative specifications yield radically different results, Twenge et al. prove that SCA is not a neutral, automated path to truth but remains subject to human judgment. The article

credibly establishes that if one accepts the premise that school and family problems are outcomes rather than causes of social media use, the association with mental health is practically significant. Whether that premise is correct, however, remains an open question that this study cannot answer.

The Bottom Line

Twenge et al. successfully demonstrate that the "tiny" association between screen time and mental health reported in previous research is not a fixed fact but a result of specific analytical choices. However, their counter-claim of a "substantial" effect depends entirely on the unproven assumption that variables like school unhappiness are caused by social media rather than being pre-existing confounders. Consequently, the true strength of the association likely lies somewhere between the negligible estimates of the original analysis and the substantial estimates presented here, with the exact magnitude remaining uncertain.

Potential Issues

Exclusion of control variables based on an unproven causal model: The article's central finding of a large association between social media use and poor mental health depends critically on the decision to exclude a set of control variables that the original study by Orben and Przybylski had included. The authors label these variables—such as "negative attitudes toward school," "closeness with parent," and "school grades"—as "potential mediators" and argue that controlling for them would lead to overadjustment bias (p. 2). This decision is justified by citing methodological literature that warns against controlling for variables that lie on the causal pathway between an exposure and an outcome. However, this rests on the strong, unproven assumption that the causal pathway is indeed: Social Media Use → School/Family Problems → Poor Mental Health. An equally plausible alternative model is that a third factor, such as a negative home environment or a pre-existing psychological trait, acts as a common cause (a confounder) that leads to both increased social media use and problems at school and home. In the confounding scenario, controlling for these variables would be the correct analytical procedure. The article does not provide empirical evidence or a formal causal model, such as a Directed Acyclic Graph, to justify its classification of these variables as mediators rather than confounders. This single, contestable decision appears to be the primary driver of the increase in the reported effect size. For example, in the Millennium Cohort Study data for girls, the median beta for social media use and mental health is -0.01 when including these controls, but it increases twentyfold to -0.20 when they are excluded (p. 4, Table 1a). This creates a potential circularity where a causal assumption is used to justify an analytical method that produces a large correlation, which is then presented as evidence supporting the initial causal assumption.

Causal framing of findings from cross-sectional data: The article frames its research using a strong causal narrative that is not fully supported by its cross-sectional data.

The introduction presents the rise in adolescent mental health issues as a "mystery" for which social media is the "prime suspect" (p. 1). The conclusion revisits this theme, stating it "seems plausible that increases in digital media use might be responsible for the increases in adolescent depression and anxiety that began around 2012" (p. 10). This framing encourages a causal interpretation of the findings. However, the data from all three studies are correlational and cannot distinguish whether social media use causes poor mental health, poor mental health leads to greater social media use (reverse causation), or unmeasured third variables cause both (confounding). While the authors are careful to use correlational language such as "linked to" and "association" when reporting results, the overarching argumentative structure of the article uses acausal evidence to advance a causal conclusion. The article does acknowledge this limitation by stating that "social media should remain on the list of possible explanations," but the dominant narrative may overstate the certainty of the evidence presented (p. 10).

Use of an invalid proxy variable for social media in one dataset: The analysis of the Youth Risk Behavior Surveillance System (YRBSS) dataset, one of the three pillars of the article's argument, relies on a proxy variable with questionable construct validity. The authors acknowledge that the YRBSS "does not have any questions that focus specifically on social media usage" (p. 6). Instead, they use a broad measure of "electronic device use," which explicitly includes "computers and gaming consoles" and was later updated to mention "smartphones and tablets" (p. 6). This variable encompasses a wide range of heterogeneous activities, such as homework, gaming, and video streaming, that are distinct from the article's primary construct of interest: social media. Despite this mismatch, the article's abstract and discussion frame the YRBSS results as consistent evidence for the harms of "social media use" specifically. The authors speculate that the observed association is driven by the unmeasured social media component of this variable, but this is an unsupported assertion (p. 8). The use of this invalid proxy weakens the claim of a consistent finding across three

large datasets.

Omission of key psychological confounding variables: The analysis does not control for or discuss major potential psychological confounders, such as pre-existing personality traits. Stable individual differences like neuroticism, low self-esteem, or high rejection sensitivity are established predictors of both a tendency to use social media more heavily and a greater vulnerability to mental health problems. The observed association could therefore be partially or wholly explained by these underlying dispositions rather than by social media use itself. While the article re-analyzes existing datasets and is constrained by the variables available within them, the absence of controls for such fundamental individual differences represents a significant limitation. This leaves a powerful alternative explanation for the core findings unexamined and may lead to an overestimation of the independent effect of social media.

Subjectivity in the application of Specification Curve Analysis: The article critiques the analytical choices of Orben and Przybylski as arbitrary and effect-attenuating, but replaces them with an alternative set of choices that could also be seen as subjective and which result in a larger effect. For instance, the authors criticize the original analysis of the Millennium Cohort Study for allowing the parent-reported Strengths and Difficulties Questionnaire (SDQ) to constitute "73% of the data using scales" (p. 3). Their proposed solution is to weight each of the four mental health scales equally. While presented as a neutral correction for representativeness, the authors also note that "The SDQ... produces notably lower betas than the 3 other mental health measures" (p. 4). Therefore, the decision to weight all scales equally is also an active choice to down-weight the influence of the measure showing the weakest association, which mechanically increases the median beta. This highlights how Specification Curve Analysis (SCA), a technique designed to mitigate "researcher degrees of freedom," is itself highly sensitive to the researcher's decisions in defining the specification space. The article uses SCA

less to characterize this profound model uncertainty and more as a tool to argue that its preferred set of specifications is correct and the original set was wrong.

Tension between individual-level analysis and a proposed network-level mechanism: The article's conclusion introduces a network spillover hypothesis that may undermine the validity of its own individual-level analytical approach. The authors suggest that as social media became ubiquitous, "teen social life changed even for adolescents who spent no or little time on social media" (p. 10). If the primary mechanism of harm is a systemic change to the social environment that affects all adolescents regardless of their personal usage levels, then an analysis correlating an individual's hours of use with their individual mental health outcome is likely to be misspecified. This approach implicitly treats low-users as a valid comparison group, but the authors' own hypothesis suggests this group is also affected by the environmental change. Consequently, the individual-level correlation measured in the article may not capture the total effect of social media on adolescent mental health and could be a biased estimate. The article acknowledges this possibility as an avenue for "future research" but does not fully grapple with how this theoretical claim challenges the foundational premise of the study's own design (p. 10).

Potential for unaddressed systematic measurement error: The study's conclusions are based on self-reported measures of technology use, which are known to be imprecise and subject to potential biases. The authors acknowledge this limitation and suggest that better measurement would likely reveal "substantially larger" correlations (p. 10). This assumes that the measurement error is random, which tends to attenuate correlations. However, it does not account for the possibility of systematic measurement error. For example, adolescents who are depressed may be more likely to recall their screen time as being excessive or may ruminate more on their social media use, leading them to over-report their usage. Such a bias would create a spurious association between mental health and reported use that is not reflective of a true causal effect. The article does not address this possibility, and its optimistic

assumption that better measurement would strengthen its findings is speculative.

Restriction to linear relationships: The analysis exclusively models a linear relationship between the duration of social media use and mental health outcomes. The article acknowledges that "It is very likely that many of the associations are curvilinear or threshold structured" (p. 10). For instance, it is plausible that light-to-moderate use has a different, perhaps even beneficial, effect compared to heavy use. By not testing for non-linear relationships, the analysis may mischaracterize the nature of the association. A linear model could average out a null or positive effect for a majority of users with a strong negative effect for a small group of very heavy users, resulting in a misleading summary of a moderate negative effect for all. The authors state that this was a practical decision to limit the complexity of the specification space, but it remains a significant limitation on the interpretation of the findings (p. 10).

Overstated narrative emphasis on a secondary analytical choice: The article's narrative places heavy emphasis on the critique that the analysis by Orben and Przybylski was flawed because one mental health scale (the SDQ) dominated the specification space. However, the article's own results in Table 1a suggest this factor was less influential than the choice of control variables (p. 4). When the "potential mediator" controls are included in the models, the median beta is near zero regardless of whether the scales are weighted equally (-0.01) or the SDQ dominates (0.01) (p. 4, Table 1a). The large difference between the weighting methods only emerges once the mediator controls are removed. This indicates that the decision to exclude controls is the dominant analytical choice driving the results, and the narrative's focus on the scale weighting issue may overstate its relative importance.

Minor presentation and clerical issues: Several minor issues in the article's presentation may affect clarity. First, the term Specification Curve Analysis (SCA) is applied to analyses with very few specifications; in the case of the Monitoring the Future dataset, the analysis is described as an "SCA... composed of a single specifi-

cation" regarding the well-being item, though it does iterate through other variables (p. 6). Second, Tables 2a and 2b include a column for a variable titled "Internet news never/every day" that is not analyzed in the text (pp. 6–7). Finally, in the YRBSS analysis, the reported beta coefficients for girls (-0.11) and boys (-0.09) remain identical to two decimal places across three different sets of control variables (p. 8, Tables 3a and 3b), a stability the authors attribute to the lack of mediator controls in the original Orben and Przybylski analysis for this specific dataset.

Future Research

Longitudinal mediation analysis: Future work should utilize longitudinal datasets with cross-lagged panel designs to empirically test the direction of causality between social media use and the disputed control variables (e.g., school unhappiness, family closeness). Determining whether these factors function primarily as mediators or confounders is essential to resolving the discrepancy between the Orben and Przybylski and Twenge et al. estimates.

Objective measurement integration: To address the limitations of self-report data and the "electronic device" proxy issues, researchers should prioritize studies linking objective digital trace data (logs of specific app usage) with mental health outcomes. This would allow for a precise distinction between social media, gaming, and general screen time, eliminating the construct validity problems found in datasets like the YRBSS.

Non-linear modeling: Future analyses should move beyond linear correlations to explicitly model threshold effects or curvilinear relationships. As the authors note, it is plausible that the relationship is harmless at low levels and deleterious only at high levels; testing for these non-linearities could reconcile the mixed findings and provide more nuanced public health guidance.

isitcredible.com