Meta-Psychology, 2021, vol 5, MP.2020.2711
https://doi.org/10.15626/MP.2020.2711
Article type: Commentary
Published under the CC-BY4.0 license

Open data: Not Applicable
Open materials: Yes
Open and reproducible analysis: Yes
Open reviews and editorial process: Yes
Preregistration: No

Edited by: Rickard Carlsson
Reviewed by: Streamlined Peer Review
Analysis reproduced by: Lucija Batinović
All supplementary files can be accessed at OSF:
https://doi.org/10.17605/OSF.IO/XWVED

Does the privacy paradox exist? Comment on Yu et al.'s (2020) meta-analysis

Tobias Dienlin
University of Vienna

Ye Sun
City University of Hong Kong

Abstract

In their meta-analysis on how privacy concerns and perceived privacy risk are related to online disclosure intention and behavior, Yu et al. (2020) conclude that "the 'privacy paradox' phenomenon (...) exists in our research model" (p. 8). In this comment, we contest this conclusion and present evidence and arguments against it. We find five areas of problems: (1) Flawed logic of hypothesis testing; (2) erroneous and implausible results; (3) questionable decision to use only the direct effect of privacy concerns on disclosure behavior as evidence in testing the privacy paradox; (4) overinterpreting results from MASEM; (5) insufficient reporting and lack of transparency. To guide future research, we offer three recommendations: Going beyond mere null hypothesis significance testing, probing alternative theoretical models, and implementing open science practices. While we value this meta-analytic effort, we caution its readers that, contrary to the authors' claim, it does not offer evidence in support of the privacy paradox.

Keywords: Privacy Paradox, Meta-Analysis, Comment

In a recent meta-analysis on how privacy concerns and perceived privacy risk are related to online disclosure intention and behavior, the authors conclude that "privacy concern cannot significantly affect disclosure behavior, which confirms that the 'privacy paradox' phenomenon [. . .] exists in our research model" (Yu et al., 2020, p. 7f.). Such a strong claim from a meta-analytic study is likely to impact future research on online privacy in substantial ways. In this comment, we challenge this conclusion and present contesting evidence and arguments. While we value this meta-analytic effort, we caution its readers that, contrary to the authors' claim, it does not offer evidence in support of the privacy paradox. We first describe and discuss five areas of problems in Yu et al.'s (2020) analysis. Based on these problems, we then offer three recommendations for future research.

Problem 1: Flawed logic of hypothesis testing

The privacy paradox phenomenon describes "the dichotomy between information privacy attitudes and actual behavior" and (Kokolakis, 2017, p. 122). For example, despite stating that they are concerned about privacy, users still share much information online.

Empirically, the privacy paradox is tested by analyzing the relationship between privacy cognitions (e.g., privacy attitudes, privacy concerns, or perceived privacy risk) and privacy behavior (e.g., information disclosure or privacy protection) (Gerber et al., 2018). In primary studies, rejecting the privacy paradox hypothesis is relatively straightforward when there is a significant, negative relationship between cognitive and behavioral variables. In other words, if increased privacy concerns are associated with reduced online sharing, such evidence refutes the privacy paradox (e.g., Utz & Krämer, 2009). "Support" for the paradox, on the other hand, is typically inferred from the lack of a significant relationship (Taddicken, 2014). To date, only a few studies have found that privacy cognitions are positively related to disclosure outcomes (e.g., Contena et al., 2015), which would constitute direct support for the privacy paradox.

Yu et al. (2020, p. 4) formally test the privacy paradox via a null hypothesis, which states, "H4: Privacy concern has no significant effects on users' personal information disclosure behavior. Namely, privacy paradox exists." The logic of this hypothesis testing is problematic, however, because absence of evidence is not evidence of absence (e.g., a sample of all white swans is no evidence that black swans do not exist). A nonsignificant result (i.e., $p > .05$) does not mean that the null hypothesis is true or should be accepted (see Greenland et al., 2016, for detailed discussions of this and related misperceptions).

In sum, a finding of no significant effects cannot demonstrate the absence of the effect (and hence the existence of the privacy paradox).

Problem 2: Erroneous and implausible results

To demonstrate the hypothesized null effect, Yu et al. (2020) show a non-significant direct effect of privacy concerns on disclosure behavior. We first demonstrate the errors in the evidence presented in their report. We then critically discuss the authors' decision to focus only on this direct effect in Problem 3 and their use of meta-analytical structural equation modeling (MASEM) in Problem 4.

The authors conclude the non-significant direct effect of the privacy concerns on disclosure behavior by comparing two structural equation models. The proposed (and final) model does not include the direct effect (see Figure 1, top panel), whereas the saturated model does (Figure 1, bottom panel). The two models are compared based on only one criterion: the RMSEA index. According to the authors, the proposed model shows a good model fit with an RMSEA = .008. For the saturated model, the authors write that "[t]he model fit indices [. . .] were not acceptable (RMSEA = 0.368). This implied that our proposed model was effective and that privacy concerns could not predict users' disclosure behavior. Thus, our H4 was supported, which indicated that privacy paradox does exist." (Yu et al., 2020, p. 5).

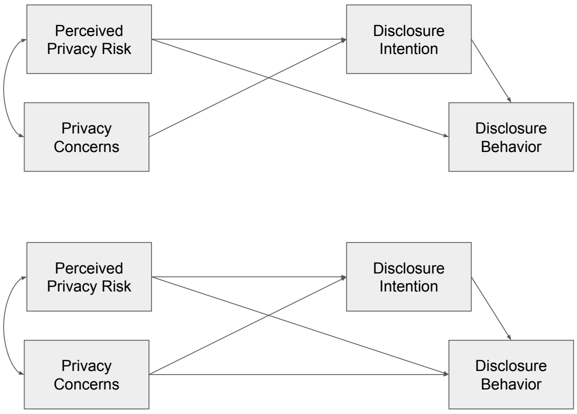

Figure 1. Top: Proposed/final model reported in Yu et al. (2020). Bottom: Saturated model to test the direct effect of privacy concern on disclosure behavior in Yu et al. (2020).

The results reported by the authors are erroneous and implausible. First, because leaving out a relationship in a path model effectively constrains it to zero (Kline, 2016), which is unlikely in most cases (e.g., Orben & Lakens, 2020), model fit should *increase* if we add another path (Kline, 2016). For the saturated model, with an added path between privacy concerns and disclosure behavior, a decreased model fit is implausible. Second, the reported RMSEA of .368 for the saturated model must be erroneous. The RMSEA for a saturated model (with $df = 0$) should be undefined (or zero), as can be seen from the formula for its calculation (Kline, 2016, p. 205).

$$RMSEA = \sqrt{\frac{\chi^2_M - df_M}{df_M(N-1)}} \tag{1}$$

We reran the model with the correlation matrix reported in Table 1 in Yu et al. (2020) (for the syntax and the results, see online supplementary material at https://osf.io/qexpf/). Following the authors' procedure, we calculated the harmonic mean using the sample statistics provided in Table 1. As expected, our reanalysis showed a "perfect" fit with an RMSEA of 0. In addition, for the proposed model, we were able to reproduce all the fit indices reported in Yu et al. (2020) except for the RMSEA: Whereas the authors report an RMSEA of .008, our re-analysis produced an RMSEA of .08. Therefore, the RMSEAs reported in Yu et al. (2020) – .368 for the saturated model and .008 for the final model – are both erroneous. These errors invalidate the findings that the authors use as the key evidence for the privacy paradox hypothesis testing.

We also quickly point out some other more minor problems we have observed. First, the bivariate correlation between privacy concerns and disclosure behavior is reported to be r = –0.063 in Table 1 but as "$r_{\text{PC-DB}}$ = –0.120" in the text (p. 6). Second, at least twice in the text, the authors consider a p-value between .05 and .10 as relevant/statistically significant, without acknowledging or explaining the shift from the usual significance level of 5%. Finally, using only RMSEA as the basis for model comparison is subpar. Besides, for models with low degrees of freedom, RMSEA is problematic and should be avoided (Kenny et al., 2015).

Problem 3: The questionable decision to use only the direct effect (or the lack thereof) of privacy concerns on disclosure behavior as evidence in testing the privacy paradox.

Notably, the authors' claim for evidence regarding the privacy paradox is *only* based on (the lack of) the *direct* effect of *privacy concerns* on disclosure behavior. On theoretical and methodological grounds, we argue that the authors' decision to conclude privacy paradox solely based on this one path is problematic.

Problem 3a. The omission of indirect effect via behavioral intention. Yu et al.'s (2020) Hypothesis 4 (see above) addresses the overall effect of privacy concerns on disclosure behavior in testing the privacy paradox, not just its residual direct effect. In presenting evidence for H4, however, the authors entirely omit the *indirect* effect of privacy concerns on disclosure behavior via disclosure intention.

Statistically, this decision is peculiar: When estimating the overall effect of the independent variable on the outcome, direct and indirect effects are to be combined. Theoretically, this omittance lacks justification as well. According to Theory of Planned Behavior (Fishbein & Ajzen, 1975), on which Yu et al.'s (2020) model is based, attitudes affect behavior indirectly via behavioral intentions.

Behavioral intentions, as a mediator, help explain how and why an effect takes place. The residual direct effect in a model often captures the influence of unexamined mechanisms. Neither the indirect or the direct effect alone addresses the existence of an effect or its magnitude (Rohrer, 2018). To this end, in testing H4, the authors should estimate the total effect.

Statistically identical to the total effect is the bivariate correlation between the two variables (Hayes, 2013), which is provided in the paper. Notably, Table 1 in Yu et al. (2020) reports a *significant* correlation between privacy concerns and disclosure behavior (r = –0.063, 95% CI [–0.120; –0.005], p = .034). Granted, it is a very small effect (see below) – but if we just use the p-value for hypothesis testing (e.g., $p < .05$), which is the approach used in the paper, the conclusion would be to reject the privacy paradox.

Problem 3b. The exclusion of privacy risk perceptions from the privacy paradox framework. By focusing on the residual direct effect of privacy concerns on disclosure behavior, the authors in effect claim that risk perceptions, treated as a confounding variable, have no theoretical or empirical role in the privacy paradox framework. We find this decision questionable.

We agree with the authors' position that privacy concerns and perceived risk are conceptually distinguishable. Nonetheless, to make such a distinction does not mean that privacy cognitions, as a larger construct in the privacy paradox, are represented *only* by privacy concerns. Instead, we argue that risk perceptions are also relevant in testing the privacy paradox. As the authors evoke in their literature review, Gerber et al. (2018, p. 245) explicitly list "privacy attitude, concerns, perceived risk, behavioral intention and behavior" as central variables for "privacy paradox explanation attempts". When providing examples of the privacy paradox in the introduction, Yu et al. (2020) themselves include research on perceived privacy risks.

In any event, making a conceptual distinction between the two variables should not lead to disregarding their close relationship. Despite their differences (see also below), both concepts capture cognitions toward privacy. Empirical data show high correlations between the two: r = .73 in (Bol et al., 2018), and r = .62 as is reported in this meta-analysis.

Therefore, theoretically and empirically, privacy concerns and risk perceptions are *both* part of the privacy paradox framework. The role of risk perception should not be excluded from the empirical evidence regarding the privacy paradox. Yu et al.'s (2020) analysis finds that perceived privacy risk is a significant predictor of online privacy behavior (r = –.165, p = .003). In other words, people who perceive online information sharing as riskier also share less information, which we believe represents compelling evidence against the privacy paradox.

Problem 4: Overinterpreting MASEM results

We also caution against the authors' overinterpretations of their MASEM results as evidence of causal relationships. The authors claim that they "conducted structural equation modeling, based on meta-analytically pooled correlations (MASEM), to investigate the *causal* effects of [. . .] privacy cognition on online disclosure intention and behavior" (Yu et al., 2020, p. 2, emphasis added).

The use of structural modeling techniques alone does not ensure causal inferences. Results of model fitting should not be interpreted as if they came from an experiment when they were not (Loehlin & Beaujean, 2016). In addition, "the data do not confirm a model, they only fail to disconfirm it" (Cliff, 1983, p. 116). There are equivalent models that also fit the data, and there are unanalyzed variables that could disconfirm the model if included. Most models in the social science research are "descriptive models" that simply depict relationships but are presented as "structural models" yielding causal explanations (Freedman, 1987, p. 221).

Yu et al.'s (2020) model, if estimated correctly, could provide descriptive relationships but not causal evidence. The included primary studies typically analyzed cross-sectional, self-reported data. The MASEM approach used in Yu et al. (2020) also does not incorporate potential confounding variables such as age, sex, or education level, which may affect the relationship between privacy cognitions and online sharing behavior (Kezer et al., 2016; Tifferet, 2019). The small number of variables and degrees of freedom in the model also limits the usefulness of MASEM (Cheung, 2021) and leaves little room for testing possible alternative models.

Yu et al.'s (2020) particular approach to MASEM is also a limited one with problematic statistical properties. MASEM includes a collection of methods that combine meta-analysis and SEM. What Yu et al. (2020) used is the *univariate-r approach*, which first meta-analyzes each correlation as if they were independent and then fits an SEM on the pooled average correlation matrix as if it were an observed covariance matrix (Cheung, 2019, 2021). Other MASEM methods, varying in specific procedures, are multivariate approaches that aggregate correlational matrices from primary studies by taking into consideration the dependence of correlations. The latter also allows for handling missing data and addressing estimation uncertainty in fitting the SEM. The univariate-r approach has known statistical issues (Cheung, 2021). For example, the pairwise aggregation/deletion means that an ad-hoc sample size is used for SEM (the harmonic mean is the most common), which leads to biased test statistics and standard errors. The SEM results also differ depending on which ad-hoc sample size is used. Ignoring sampling uncertainty across studies and treating the correlation matrix as the covariance matrix have also been shown to generate incorrect estimates (Cheung, 2019, 2021).

Overall, we remind readers that Yu et al.'s (2020) MASEM results should be interpreted with great caution. The use of a structural model does not automatically allow for causal inferences, and the implemented univariate-r approach to MASEM leads to estimates that are, in general, less trustworthy.

Problem 5: Insufficient reporting and lack of transparency

Finally, there is a substantial underreporting in Yu et al.'s (2020) meta-analysis. Meta-analyses, like other empirical research work, are under the influence of researchers' subjective decisions and common errors. Standardized reporting and data transparency are important to the reproducibility of meta-analytic findings (Lakens et al., 2016; Maassen et al., 2020). Yu et al.'s (2020) reporting does not adhere to the reporting guidelines for meta-analysis such as the PRISMA statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) (Moher et al., 2009). Key information is not available in either the published paper or the supplemental material. For example, individual effect estimates are not reported. The inclusion and exclusion criteria are described only vaguely. There is no description about how the key variables (such as privacy concerns vs. perceived privacy risks) were operationalized and extracted from the primary studies. Regarding moderator coding, no information was provided about coder training or inter-coder reliability. No publication bias assessment was reported.

The lack of transparency in Yu et al. (2020) makes it hard to assess the validity of their reported meta-analytic data. This comment and the additional analyses we report are consequently constrained despite our best efforts. Without sufficient information to evaluate the search, coding, or effect size extraction processes in Yu et al.'s (2020) meta-analysis, nor any data to reproduce their summary effect sizes, we remind readers that re-analysis can only be as good as the available data it is based on, the quality of which we are unable to assess.

Future Research

The five major areas of problems with Yu et al.'s (2020) analysis, as discussed above, put their conclusion regarding the privacy paradox in question. These problems, to a certain extent, speak to larger issues and challenges in the existing privacy paradox research, meriting further reflections. In this section, we engage in such reflections and provide three recommendations for future research on the privacy paradox.

Recommendation 1: Going beyond null hypothesis significance testing (NHST)

As we discussed in Problem 1, to "confirm" a null hypothesis via statistical non-significance is a flawed approach. This misconception of NHST is not specific to Yu et al. (2020). For example, in examining the privacy paradox, Hallam and Zanella (2017) also proposed a null hypothesis that “privacy concern is not related to sensitive information self-disclosure behavior” (p. 220) and used the non-significant effect as supportive evidence. A *p*-value above .05 alone cannot distinguish a true null effect from insensitive data (Dienes, 2014; Lakens et al., 2018).

By contrast, rather than relying on the *p*-value as the sole arbiter of “truth” in NHST, to assess whether there is evidence for the null effect researchers can adopt an *interval perspective* or calculate *Bayes factors*. In what follows, we introduce and illustrate these two approaches using the data from Yu et al. (2020) and from Baruh et al. (2017), which is another meta-analysis on the privacy paradox. The results are to be read in the context of the example and not as a definitive answer to the existence of the privacy paradox.

1a. The interval estimates approach

Interval estimates include Bayesian credibility intervals or frequentist confidence intervals, which are “the set of possible population values consistent with the data; all other population values may be rejected” (Dienes, 2014, p. 3). An interval approach can overcome the problems associated with NHST by determining a *range* of values consistent with a hypothesized effect. The null hypothesis, therefore, no longer hinges on a single point value (such as 0 in most null hypotheses in NHST) but specifies a “null region” (Dienes, 2014).

Delineating the null region requires determining a minimally interesting effect size, or the so-called “smallest effect size of interest” (SESOI, Lakens et al., 2018). We can then make statistical inferences by comparing the observed interval against the null region, following the guidelines outlined in Dienes (2014). For example, if an effect is too small to be meaningful, a hypothesis is rejected – even if the *p*-value is below 5%. Using a null region with a pre-defined SESOI, researchers can therefore better assess the empirical evidence in terms of its actual theoretical and practical significance.

We illustrate below how these rules may apply in the context of the privacy paradox research. We set a predetermined SESOI of r = -.05 (Funder & Ozer, 2019), hence a null region of r = -.05 to .05, and depict hypothetical interval estimates in the upper panel of Figure 2. Corresponding to the four rules in Dienes (2014), these depicted scenarios are interpreted as follows:

1. *Rule 1: If the interval falls completely within the null region, accept the null region hypothesis*. This case, therefore, presents evidence that leads to the acceptance of the privacy paradox hypothesis (i.e., privacy concerns are unrelated to disclosure behavior).

2. *Rule 2: If the interval falls completely outside of the null region, reject the null region hypothesis*. In this case, as the interval has no overlap with and is on the negative side of the null region, it is unambiguous evidence for the alternative hypothesis (i.e., privacy concerns are negatively related to disclosure behavior). The privacy paradox is rejected.

3. *Rule 3: If the interval overlaps with the null region only on one side, reject the directional hypothesis accordingly*. In this case, the upper limit of the interval is below the SESOI, thereby rejecting the positive effect hypothesis. In other words, the hypothesis that there is a positive relationship between privacy concerns and disclosure behavior (i.e., a stronger version of the privacy paradox) is rejected. We suspend judgement regarding a negative effect hypothesis or a null region hypothesis.

4. *Rule 4: If the interval contains values both above and below the null region, suspend judgement*. In this case, the observed interval overlaps with the null region beyond both sides, which means that the data are insensitive; thus, no conclusion can be made.

The lower panel of Figure 2 displays real data from the two meta-analyses on the relationship between privacy concerns and information sharing: the 95% confidence intervals of the overall effect sizes. Baruh et al.’s (2017) data ([-.18, -.07]) fall entirely outside and below the null region ([-.05, .05]), thus squarely rejecting the privacy paradox. To interpret Yu et al.’s (2020) data ([-.01, -.12]), first, the null region hypothesis cannot be accepted (i.e., Rule 1 does not apply). Second, the positive directional hypothesis is rejected (i.e., Rule 3 applies), meaning that there is evidence against a positive relationship between privacy concerns and disclosure behavior. We would suspend judgment regarding the existence of a negative effect or no effect.

1b. The Bayes factor approach

Another alternative is to use Bayes factors, which compare the probability of two competing hypotheses – in this case, an alternative hypothesis and the null hypothesis (Dienes, 2014). Bayes factors (B), ranging from 0 to infinity, indicate that “data are B times more likely under the alternative than under the null” (Dienes, 2014, p. 4). A value of B larger than 1, therefore, suggests greater evidence for the alternative hypothesis than the null. Although there is no absolute cutoff point for B (unlike the dichotomized *p*-value for statistical significance), the following guideline has

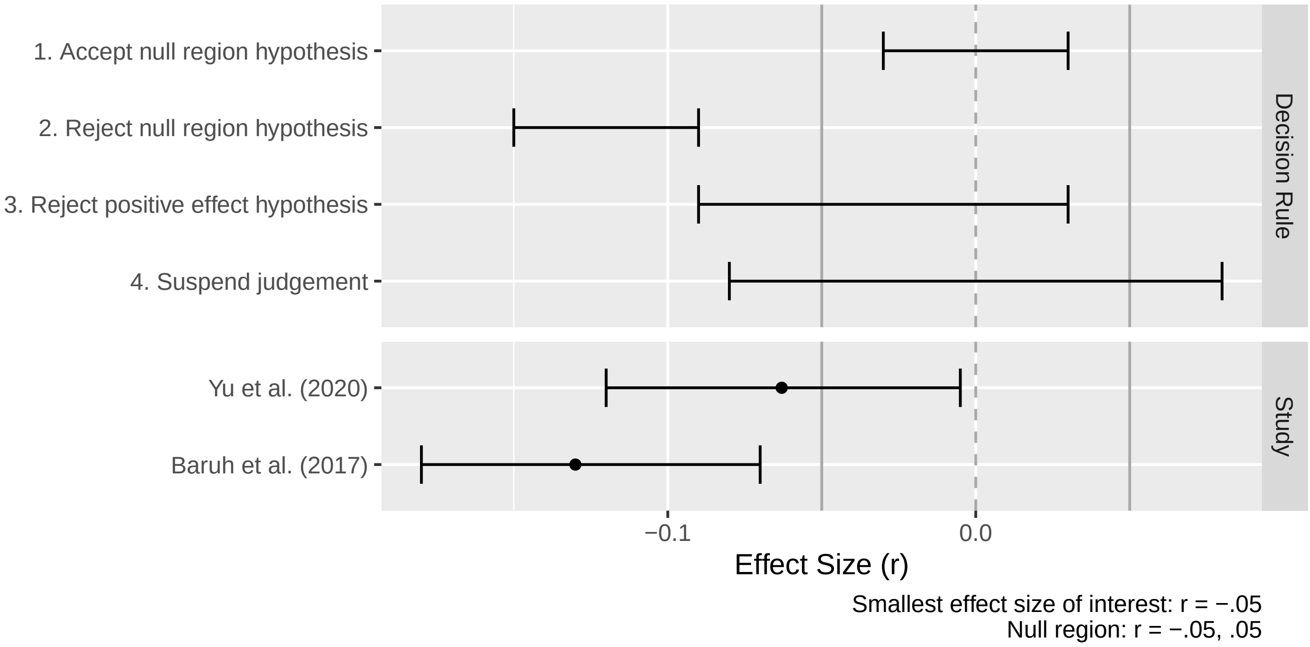

Figure 2. Upper panel: Illustration of the decision rules proposed by Dienes (2014) using hypothetical data. Lower panel: 95% confidence intervals of the relation between privacy concerns and information sharing as reported by the two meta-analyses: Yu et al. (2020) & Baruh et al. (2017). Conclusion: Baruh et al.'s (2017) data reject the null region (i.e., the privacy paradox) hypothesis. For Yu et al.'s (2020) data, the positive effect (i.e., privacy concerns increase disclosure behavior) hypothesis is rejected.

been suggested to ease its interpretation: A *B* value greater than 3 indicates "substantial" (Jeffreys, 1961) or, more recently, "moderate" (Lee & Wagenmakers, 2013) evidence for the alternative hypothesis; a *B* of lower than 1/3 indicates substantial/moderate evidence for the null hypothesis; and the values in-between are considered weak or anecdotal evidence (Dienes, 2014).

To apply Bayes factors to the privacy paradox research, we could postulate a small negative correlation (r = -.10) for the alternative hypothesis and a null effect (r = 0) for the privacy paradox hypothesis. We then compare the two hypotheses by calculating a Bayes factor (Dienes, 2008), assuming that the effect is normally distributed with a standard deviation of r / 2 = .05 (see Dienes, 2014). Using the data from Yu et al. (2020), the resulting *B* is 3.96 (see online supplementary material at https://osf.io/qexpf/). In other words, the alternative hypothesis of a small negative effect is about four times more likely than the null hypothesis, which constitutes at least moderate evidence against the privacy paradox.

Similarly, instead of referring to a null effect, using Bayes we can also compare other informative hypotheses. For example, combining Bayes and the SESOI logic, we can compare the probability of all meaningful negative effects (say, H1: r < -.05) versus its complement (i.e., H2: r ≥ -.05). In this case, H2 therefore captures all values we consider "paradoxical", including null effects and positive effects. Using the R package "bain" (Van Lissa et al., 2020) and the data from Yu et al. (2020), we compared both hypotheses and found that H1 is 7.54 times more likely than its complement H2. In other words, with the data collected from Yu et al. (2020), the theory that there is no privacy paradox would be about 7 times more likely than the theory that the privacy paradox exists.

Recommendation 2: Rethinking the theoretical/conceptual model

As we discussed in Problem 3, the theoretical relationship between privacy concerns and risk perceptions is open to rethinking and subject to future empirical testing. In Yu et al.'s (2020) proposed/final model, privacy concerns and perceived privacy risks are modeled as parallel predictors of disclosure intentions and/or behavior, and perceived risks are treated as a control variable (see top panel of Figure 1). We encourage future research to consider alternative models. For example, perceived privacy risks can be posited as a *mediator*

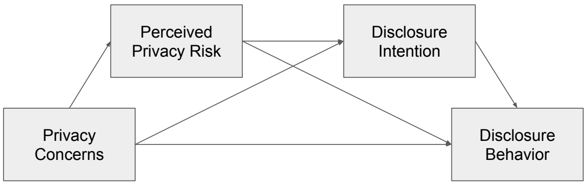

Figure 3. Our suggested theoretical model for future research, in which perceived privacy risks and disclosure intention mediate the effect of privacy concerns on disclosure behavior.

between privacy concerns and disclosure behavior (see Figure 3).

To explain, privacy concerns are often conceptualized as general, trait-like, and intuitive factors; perceived privacy risks, by contrast, are often understood as specific, state-like, and rational factors (Bol et al., 2018). Because general dispositions often precede more specific cognitions (Fishbein & Ajzen, 1975), general concerns about privacy may likewise shape more specific perceived privacy risk (Dienlin & Trepte, 2015; Heirman et al., 2013). Supporting our theoretical model, a large body of empirical research also analyzed privacy concerns as predictors of perceived privacy risk (e.g., Keith et al., 2013; Lancelot Miltgen et al., 2013; Li et al., 2011; Zhou, 2015, see the review in Gerber et al. 2018).

We encourage future researchers to take up the theoretical and empirical tasks of explicitly clarifying the relationship between privacy concerns and risk perceptions. Finding the "correct" model is important. Statistically controlling for variables that really are mediators will lead to false results (Rohrer, 2018). Such clarifications will enable more precise modeling and hence more accurate evidence in examining the privacy paradox.

Recommendation 3: Implementing open science practices

Researchers are humans, and humans make mistakes. Reporting errors such as numerical inconsistencies are common in social sciences in general (see the review in Nuijten et al., 2016). Whereas human errors are often inevitable, we as researchers should nonetheless help one another enhance the rigor of our research processes to avoid, detect, and correct errors.

To ensure a discipline's self-scrutiny and hence self-correction, we encourage online privacy researchers to increase openness and transparency. The recent open science movement in social sciences (Munafò et al., 2017; Nosek et al., 2015), evoking Mertonian norms such as communalism and organized skepticism (Merton, 1942), seeks to promote such values and norms in research practices. Transparency is key to improving research reproducibility and replicability, a major goal in the face of the replication crisis (Camerer et al., 2016; Camerer et al., 2018; Open Science Collaboration, 2015). Open science practices include adherence to reporting standards in publications, preregistration of study plans, data sharing, and reproducible workflow documentation (for overviews, see Christensen et al., 2019; Dienlin et al., 2021; Munafò et al., 2017). Transparency also means greater efficiency for the research community, as we can share resources for error-checking, replication, and developing new studies.

For future meta-analyses, we encourage researchers to engage more open science practices to build cumulative research. We specifically recommend pre-registering analyses, complying with the reporting standards, and making data and other essential materials of the research process publicly available (Lakens et al., 2016).

Conclusion

Meta-analyses do not offer conclusive findings for an area of research. Notably, and not discussed in Yu et al. (2020), another meta-analysis on the privacy paradox finds a negative significant relationship between privacy concerns and information sharing ($r = -.13$; Baruh et al., 2017), which speaks against the privacy paradox. The meta-review by Gerber et al. (2018, p. 226) concludes that "[...] *strong predictors* for privacy behavior are privacy intention, willingness to disclose, *privacy concerns* and privacy attitude" (emphasis added). This meta-analysis by Yu et al. (2020), like others, represents only one assessment of the area of research and shall not be taken as definitive.

In this comment, we lay out evidence and arguments that question the validity of Yu et al.'s (2020) data, analyses, and results. Based on the accessible information, we contest their conclusion that the privacy paradox exists. Re-analyzing their data rather seems to provide some evidence against this paradox. As we have emphasized, due to the underreporting in Yu et al.'s (2020) paper, we are unable to assess the validity of the reported data when performing our re-analyses. This caveat, we hope, serves both to mark the limitations of this comment and to accentuate the importance of standard reporting and data transparency for empirical researchers, including meta-analysts.

In closing, we believe that the privacy paradox remains an open question in need of further theoretical and empirical efforts. We hope that this comment presents a constructive engagement with Yu et al.'s

(2020) meta-analysis and inspires more theory-based, rigorous, and open research on the privacy paradox in the future.

Author Contact

Tobias Dienlin, University of Vienna, Department of Communication, 1090 Vienna, Austria. E-mail: tobias.dienlin@univie.ac.at. ORCID: 0000-0002-6875-8083.

Ye Sun, City University of Hong Kong, Department of Media and Communication, 83 Tat Chee Avenue, Kowloon Tong, Kowloon, Hong Kong. E-Mail: yesun27@cityu.edu.hk. ORCID: 0000-0001-8551-2037.

Corresponding author: Tobias Dienlin

Acknowledgements

TD has already published on the privacy paradox, and in most of his studies he found support against the privacy paradox. We would like to thank Malte Elson, Niklas Johannes, Philipp Masur, Julia Rohrer, and Michael Scharkow for valuable feedback on this submission

Conflict of Interest and Funding

Both authors declare no conflicts of interest. While working on the manuscript, TD received funding from the Volkswagen Foundation.

Author Contributions

TD and YS wrote the article; TD ran the reanalysis and wrote the code; TD supervised the project. Authorship order was determined by magnitude of contribution.

Open Science Practices

This article earned the Open Materials badge for making the materials openly available. It has been verified that the analysis reproduced the results presented in the article. This is a commentary that analyzed a published article, and as such has no new data. The editorial process for this article relied on streamlined peer review where peer reviews obtained from previous journal(s) were moved forward and used as the basis for the editorial decision. These reviews are shared in the supplementary files, as part of the authors' cover letter. The identities of the reviewers are shown or hidden in accordance with the policy of the journal that originally obtained them. The entire editorial process is published in the online supplement.

References

Baruh, L., Secinti, E., & Cemalcilar, Z. (2017). Online privacy concerns and privacy management: A meta-analytical review. *Journal of Communication*, *67*(1), 26–53. https://doi.org/10.1111/jcom.12276

Bol, N., Dienlin, T., Kruikemeier, S., Sax, M., Boerman, S. C., Strycharz, J., Helberger, N., & Vreese, C. H. (2018). Understanding the effects of personalization as a privacy calculus: Analyzing self-disclosure across health, news, and commerce contexts. *Journal of Computer-Mediated Communication*, *23*(6), 370–388. https://doi.org/10.1093/jcmc/zmy020

Camerer, C. F., Dreber, A., Forsell, E., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Almenberg, J., Altmejd, A., Chan, T., Heikensten, E., Holzmeister, F., Imai, T., Isaksson, S., Nave, G., Pfeiffer, T., Razen, M., & Wu, H. (2016). Evaluating replicability of laboratory experiments in economics. *Science*, *351*(6280), 1433–6. https://doi.org/10.1126/science.aaf0918

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., Kirchler, M., Nave, G., Nosek, B. A., Pfeiffer, T., Altmejd, A., Buttrick, N., Chan, T., Chen, Y., Forsell, E., Gampa, A., Heikensten, E., Hummer, L., Imai, T., . . . Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, *2*(9), 637–644. https://doi.org/10.1038/s41562-018-0399-z

Cheung, M. W.-L. (2019). Some reflections on combining meta-analysis and structural equation modeling. *Research Synthesis Methods*, *10*(1), 15–22. https://doi.org/10.1002/jrsm.1321

Cheung, M. W.-L. (2021). Meta-analytic structural equation modeling. *Oxford Research Encyclopedia of Business and Management*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190224851.013.225

Christensen, G., Freese, J., & Miguel, E. (2019). *Transparent and reproducible social science research: How to do open science.*

Cliff, N. (1983). Some cautions concerning the application of causal modeling methods. *Multivariate Behavioral Research*, *18*(1), 115–126. https://doi.org/10.1207/s15327906mbr1801_7

Contena, B., Loscalzo, Y., & Taddei, S. (2015). Surfing on social network sites. *Computers in Human Behavior*, *49*, 30–37. https://doi.org/10.1016/j.chb.2015.02.042

Dienes, Z. (2014). Using Bayes to get the most out of non-significant results. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.00781

Dienlin, T., Johannes, N., Bowman, N. D., Masur, P. K., Engesser, S., Kümpel, A. S., Lukito, J., Bier, L. M., Zhang, R., Johnson, B. K., Huskey, R., Schneider, F. M., Breuer, J., Parry, D. A., Vermeulen, I., Fisher, J. T., Banks, J., Weber, R., Ellis, D. A., . . . de Vreese, C. (2021). An agenda for open science in Communication. *Journal of Communication*, *71*(1), 1–26. https://doi.org/10.1093/joc/jqz052

Dienlin, T., & Trepte, S. (2015). Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors [00201]. *European Journal of Social Psychology*, *45*(3), 285–297. https://doi.org/10.1002/ejsp.2049

Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Addison-Wesley.

Freedman, D. A. (1987). A rejoinder on models, metaphors, and fables. *Journal of Educational Statistics*, *12*(2), 206–223. https://doi.org/10.2307/1164900

Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense [00014]. *Advances in Methods and Practices in Psychological Science*, *2*(2), 156–168. https://doi.org/10.1177/2515245919847202

Gerber, N., Gerber, P., & Volkamer, M. (2018). Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior [00017]. *Computers & Security*, *77*, 226–261. https://doi.org/10.1016/j.cose.2018.04.002

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., & Altman, D. G. (2016). Statistical tests, P values, confidence intervals, and power: A guide to misinterpretations. *European Journal of Epidemiology*, *31*(4), 337–350. https://doi.org/10.1007/s10654-016-0149-3

Hallam, C., & Zanella, G. (2017). Online self-disclosure: The privacy paradox explained as a temporally discounted balance between concerns and rewards. *Computers in Human Behavior*, *68*, 217–227. https://doi.org/10.1016/j.chb.2016.11.033

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. Guilford Press. http://lib.myilibrary.com/detail.asp?id=480011

Heirman, W., Walrave, M., & Ponnet, K. (2013). Predicting adolescents' disclosure of personal information in exchange for commercial incentives: An application of an extended theory of planned behavior. *Cyberpsychology, Behavior, and Social Networking*, *16*(2), 81–87. https://doi.org/10.1089/cyber.2012.0041

Jeffreys, H. (1961). *The theory of probability* (3rd ed.). Oxford University Press.

Keith, M. J., Thompson, S. C., Hale, J., Lowry, P. B., & Greer, C. (2013). Information disclosure on mobile devices: Re-examining privacy calculus with actual user behavior [00208]. *International Journal of Human-Computer Studies*, *71*(12), 1163–1173. https://doi.org/10.1016/j.ijhcs.2013.08.016

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, *44*(3), 486–507. https://doi.org/10.1177/0049124114543236

Kezer, M., Sevi, B., Cemalcilar, Z., & Baruh, L. (2016). Age differences in privacy attitudes, literacy and privacy management on Facebook. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *10*(1). https://doi.org/10.5817/CP2016-1-2

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, *64*, 122–134. https://doi.org/10.1016/j.cose.2015.07.002

Lakens, D., Hilgard, J., & Staaks, J. (2016). On the reproducibility of meta-analyses: Six practical recommendations. *BMC Psychology*, *4*(1), 24. https://doi.org/10.1186/s40359-016-0126-3

Lakens, D., Scheel, A. M., & Isager, P. M. (2018). Equivalence testing for psychological research: A tutorial. *Advances in Methods and Practices in Psychological Science*, *1*(2), 259–269. https://doi.org/10.1177/2515245918770963

Lancelot Miltgen, C., Popovič, A., & Oliveira, T. (2013). Determinants of end-user acceptance of biometrics: Integrating the "Big 3" of technology acceptance with privacy context. *Decision Support Systems*, *56*, 103–114. https://doi.org/10.1016/j.dss.2013.05.010

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. Cambridge University Press.

Li, H., Sarathy, R., & Xu, H. (2011). The role of affect and cognition on online consumers' decision to disclose personal information to unfamiliar online vendors [00313]. *Decision Support Systems*, *51*(3), 434–445. https://doi.org/10.1016/j.dss.2011.01.017

Loehlin, J. C., & Beaujean, A. A. (2016). *Latent variable models - an introduction to factor, path, and structural equation analysis* (5th ed.). Routledge.

Maassen, E., Assen, M. A. L. M. v., Nuijten, M. B., Olsson-Collentine, A., & Wicherts, J. M. (2020). Reproducibility of individual effect sizes in meta-analyses in psychology. *PLOS ONE*, *15*(5), e0233107. https://doi.org/10.1371/journal.pone.0233107

Merton, R. (1942). A note on science and democracy. *Journal of Legal and Political Sociology*, 115–126.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, T. P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Du Percie Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1). https://doi.org/10.1038/s41562-016-0021

Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., . . . Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374

Nuijten, M. B., Hartgerink, C. H. J., van Assen, M. A. L. M., Epskamp, S., & Wicherts, J. M. (2016). The prevalence of statistical reporting errors in psychology (1985-2013). *Behavior research methods*, *48*(4), 1205–1226. https://doi.org/10.3758/s13428-015-0664-2

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), 4716. https://doi.org/10.1126/science.aac4716

Orben, A., & Lakens, D. (2020). Crud (re)defined. *Advances in Methods and Practices in Psychological Science*, *3*(2), 238–247. https://doi.org/10.1177/2515245920917961

Rohrer, J. M. (2018). Thinking clearly about correlations and causation: Graphical causal models for observational data. *Advances in Methods and Practices in Psychological Science*, *24*(2), 251524591774562. https://doi.org/10.1177/2515245917745629

Taddicken, M. (2014). The 'privacy paradox' in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication*, *19*(2), 248–273. https://doi.org/10.1111/jcc4.12052

Tifferet, S. (2019). Gender differences in privacy tendencies on social network sites: A meta-analysis. *Computers in Human Behavior*, *93*, 1–12. https://doi.org/10.1016/j.chb.2018.11.046

Utz, S., & Krämer, N. C. (2009). The privacy paradox on social network sites revisited: The role of individual characteristics and group norms. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, *3*(2). www.cyberpsychology.eu/view.php?cisloclanku=2009111001&article=2

Van Lissa, C. J., Gu, X., Mulder, J., Rosseel, Y., Van Zundert, C., & Hoijtink, H. (2020). Teacher's corner: Evaluating informative hypotheses using the Bayes factor in structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 1–10. https://doi.org/10.1080/10705511.2020.1745644

Yu, L., Li, H., He, W., Wang, F.-K., & Jiao, S. (2020). A meta-analysis to explore privacy cognition and information disclosure of internet users. *International Journal of Information Management*, *51*, 102015. https://doi.org/10.1016/j.ijinfomgt.2019.09.011

Zhou, T. (2015). Understanding user adoption of location-based services from a dual perspective of enablers and inhibitors. *Information Systems Frontiers*, *17*(2), 413–422. https://doi.org/10.1007/s10796-013-9413-1