

PERSPECTIVE/OPINION

Quality Uncertainty Erodes Trust in Science

Simine Vazire

When consumers of science (readers and reviewers) lack relevant details about the study design, data, and analyses, they cannot adequately evaluate the strength of a scientific study. Lack of transparency is common in science, and is encouraged by journals that place more emphasis on the aesthetic appeal of a manuscript than the robustness of its scientific claims. In doing this, journals are implicitly encouraging authors to do whatever it takes to obtain eye-catching results. To achieve this, researchers can use common research practices that beautify results at the expense of the robustness of those results (e.g., p-hacking). The problem is not engaging in these practices, but failing to disclose them. A car whose carburetor is duct-taped to the rest of the car might work perfectly fine, but the buyer has a right to know about the duct-taping. Without high levels of transparency in scientific publications, consumers of scientific manuscripts are in a similar position as buyers of used cars – they cannot reliably tell the difference between lemons and high quality findings. This phenomenon – quality uncertainty – has been shown to erode trust in economic markets, such as the used car market. The same problem threatens to erode trust in science. The solution is to increase transparency and give consumers of scientific research the information they need to accurately evaluate research. Transparency would also encourage researchers to be more careful in how they conduct their studies and write up their results. To make this happen, we must tie journals' reputations to their practices regarding transparency. Reviewers hold a great deal of power to make this happen, by demanding the transparency needed to rigorously evaluate scientific manuscripts. The public expects transparency from science, and appropriately so – we should be held to a higher standard than used car salespeople.

Keywords: transparency; open science; replicability; scientific integrity

In any market, consumers must evaluate the quality of products and decide their willingness to pay based on their evaluation. In science, consumers of new scientific findings must likewise evaluate the strength of the findings and decide their willingness to put stock in them. In both kinds of markets, the inability to make informed and accurate evaluations of quality (i.e., quality uncertainty) leads to a lower and lower willingness to put stock in any product – a lack of trust in the market itself. When there are asymmetries in the information that the seller and the buyer have, the buyers cannot be certain about the quality of the products, leading to quality uncertainty.

In science, quality uncertainty threatens people's ability to have confidence in findings and build on them. Here I argue that the lack of transparency in science has led to quality uncertainty, and that this threatens to erode trust in science. The solution is to require greater transparency in scientific reporting, which will increase the certainty with which quality can be evaluated, and restore trust in science.

In his paper, "The Market for "Lemons": Quality Uncertainty and the Market Mechanism" (1970), Nobel-Prize-winning economist George Akerlof illustrates this dynamic with the used car market. In this market, the seller has much more information than the buyer, making the buyer uncertain about the quality of any individual car, and thus unwilling to pay much for used cars. At extreme levels of quality uncertainty, the result is that no one is willing to buy a used car at any price – people lose all trust in the market.

There is a parallel with scientific products. In this case, the product is the manuscript or journal article, the seller is the author, and the buyer can be the journal editor, reviewers, or readers of the article – anyone who is choosing whether or not to buy the findings. The source of quality uncertainty in this market is that the authors know much more about what went into the article than do the potential buyers. There is critical information that only the authors know, including: (1) what the raw data look like, (2) what the authors' original intentions and predictions were, (3) how many studies were attempted and how many unsuccessful studies were excluded from the manuscript, and (4) how many analyses were attempted and what modifications were made before the authors settled on the analyses presented in the manuscript.

The math behind the effect of quality uncertainty on trust is quite simple. If sellers can get away with selling low quality products as if they were high quality (because buyers lack the information to tell the difference), the average quality of the products goes down. Buyers' willingness to pay is influenced by the average quality of the product in the market. Therefore, when there is a good deal of quality uncertainty, average quality will go down, driving down buyers' willingness to pay (i.e., trust in the market).

Without high levels of transparency in scientific publications, journal editors, reviewers, and readers of scientific manuscripts are in a similar position as buyers of used cars – they cannot reliably tell the difference between lemons and high quality findings. Of course the method and results sections of scientific manuscripts contain some information about the quality of the manuscript, just like the outward appearance of used cars contain some information about their quality. However, by keeping vital information private – the raw data, the original design and analysis plan, the exploratory analyses that were conducted along the way to the final analysis – authors are hiding valuable information and preventing consumers of their manuscript from being certain about its quality. Just like sellers of used cars keep things like the history of the car and any deep structural problems hidden from the buyers.

Lack of transparency is widespread in science. Several studies have shown that open data is a rare practice among researchers (Freese, 2007; Reidpath & Allotey, 2001; Vanpaemel et al., 2015; Wicherts, Boorsboom, Kats, & Molenaar, 2006). This lack of transparency is not necessarily due to deviousness or stubbornness on the part of authors – it is also encouraged by journals that place more emphasis on the aesthetic appeal of a manuscript rather than the robustness of its scientific claims (Giner-Sorolla, 2012). In doing this, journals are implicitly encouraging authors to do whatever it takes to obtain eye-catching results.

We also know, thanks to recent developments in research methods and meta-science, that researchers have many tools at their disposal to give journals what they want – overly polished results that exaggerate the quality of the research product. Many common research practices will beautify results at the expense of the robustness of those results (e.g., 'p-hacking', questionable research practices, cherry-picking, etc.; Simmons, Nelson, & Simonsohn, 2011). Some of these practices may have only small impacts on the robustness of the results (painting over a scratch on the car's body), whereas others may be glossing over deeper structural problems (a fan belt that is about to snap). Thus, the problem is not engaging in these practices, but failing to disclose them. A car whose carburetor is duct-taped to the rest of the car might still work perfectly fine,¹ but the buyer has a right to know about the duct-taping. Likewise, when readers and reviewers lack relevant details about the study design, data, and analyses, they cannot adequately evaluate the strength of a study.

This information asymmetry between authors and readers of scientific products has led to quality uncertainty, and driven the average quality of scientific papers down, to the point where there is now widespread doubt about the robustness of most of the scientific literature (Button et al., 2013; Ioannidis, 2005). According to a recent article in *Environmental Engineering Science*, "If a critical mass of scientists become untrustworthy, a tipping point is possible in which the scientific enterprise itself becomes inherently corrupt and public trust is lost, risking a new dark age with devastating consequences to humanity" (Edwards & Roy, in press).

The loss of public trust is not the only cost of quality uncertainty. In his paper, Akerlof writes "The cost of dishonesty, therefore, lies not only in the amount by which the purchaser is cheated; the cost also must include the loss incurred from driving legitimate businesses out of business." (p. 495). Again, there are parallels in science. The cost of lack of transparency is not only that we end up investing in low quality findings, and building a science on shaky foundations (which is already a significant cost), but also that we are driving rigorous science out of the market. If researchers can achieve the same result by doing shoddy science (which is cheaper, faster, and easier than rigorous science; Bakker, van Dijk, & Wicherts, 2012), there is little external incentive for them to do things the right way (Smaldino & McElreath, 2016; Tullett, 2015). When the shoddy findings later turn out not to stand up, it is too late – the high quality research has already been driven out.

Since Akerlof's article was published in 1970, we have found ways to reduce quality uncertainty in the used car market. For example, companies like Carfax® offer to uncover the information that a seller may be hiding from buyers (e.g., prior accidents). Such forensic tests may be applicable in science as well. For example, new techniques allow readers to test the integrity of the statistics presented in a manuscript (Brown & Heathers, 2016; Epskamp & Nuijten, 2015), and for larger bodies of work, meta-scientific techniques can provide information about the integrity of a set of studies (e.g., Simonsohn, Nelson, & Simmons, 2014; van Assen, van Aert, & Wicherts, 2015). Like CarFax®, these forensic tools are useful in the absence of transparency, but they are no substitute for it.

Akerlof also presents several potential solutions to the problem of quality uncertainty. For example, he suggests that intermediaries with the necessary expertise could build a business around evaluating the quality of products (i.e., a Consumer Reports-type seal of approval). This does not work for science, because one of the hallmarks of science is that the origins of scientific claims must be "available to other scholars to rigorously evaluate" (Lupia & Elman, 2014, p. 20). Indeed, the motto of the Royal Society, founded in 1660, is "take no one's word." As such, we cannot rely on a few experts to evaluate the claims, and then ask everyone else to take their word for it. Thus, we need a way to reduce quality uncertainty not only on the part of a few experts or gatekeepers (i.e., editors and reviewers), but for all who wish to read the article and make a judgment of its quality.

In the domain of used cars, and consumer products more generally, there may be no silver bullet for the problem of quality uncertainty. Happily, in science there is. Transparency would increase our certainty in the quality of scientific products tremendously. What does transparency mean? The default should be that scientists make, at a minimum, the following available:

1. All raw data on which scientific claims are based.
2. All study design plans, analysis plans, and predictions made prior to data collection (if none were made, this should be declared).
3. All relevant studies attempted and their outcomes.
4. All data exclusions and transformations.
5. All alternative models/analyses that were tested (e.g., robustness checks).

Making this the default would by no means require that all manuscripts adhere to these standards – exceptions are not only warranted, but necessary. Indeed, there are many nuances and exceptions that I am glossing over. There is, of course, the risk that we will go too far in the other direction and impose so many burdens on scientists to make everything transparent that science will grind to a halt (Lewandowsky & Bishop, 2016). However, we are nowhere near having that problem. The much more urgent threat is that we will not move fast enough towards increased transparency, and we will find ourselves with a lot full of lemons and no one interested in buying them.

How will transparency help restore trust in science? First, transparency would give ‘buyers’ the information they need to detect many misrepresentations or errors in the article. Second, the fact that buyers could potentially detect many misrepresentations would make ‘sellers’ (i.e., authors) much more accountable, and would likely increase the care with which authors conduct their studies and write up their results. By eliminating (or at least greatly reducing) information asymmetries, authors can no longer count on their errors going unnoticed. Even if such errors and misrepresentations have always been unintentional, careless mistakes, the increased accountability will motivate authors to be more careful. To the extent that any of the misrepresentations were intentional (i.e., the authors were aware that they were hiding information that would be useful to readers), this behavior should also be curbed by increased transparency, because of the increased chance of getting caught.

Of course, if a researcher is willing to manipulate the background information (e.g., fabricate raw data, or lie about their a priori plans), transparency may not always help buyers catch this. Indeed, this kind of fraud is the biggest threat to trust in science, and to the integrity of scientific findings, and we must also tackle this problem. However, solving the problem of unintentional misrepresentations would be a major advance for the scientific process, and for rebuilding trust in science.

If transparency will lead to more robust science, scientific journals (and other gatekeepers) have an obligation to insist on transparency. Unlike car dealers, journals have

a duty to serve their scientific communities, not their own bottom line. When journals choose to maximize citation impact rather than robustness, they are encouraging authors to sell them shiny- but-low-quality products, neglecting their duty. This seems to be quite common, perhaps because there are few penalties for this behavior – we continue to tie journals’ reputations to how eye-catching their articles are (i.e., impact factors). To change this, we must tie journals’ reputations to the actual quality of their articles instead (Fraleley & Vazire, 2014), and to their policies regarding transparency and openness (e.g., using the Transparency and Openness Promotion (TOP) guidelines; Nosek et al., 2015).

Top-down change is rare, and the most successful academic journals are unlikely to change on their own, for fear of harming their reputations (though see Kidwell et al., 2016 and Piwoward, Day, & Fridsma, 2007, for evidence that transparency does not seem to have a negative impact on traditional metrics, either at the journal level or the article level). Waiting for top journals to change may be a bit like hoping for profitable used car dealerships to voluntarily stop selling lemons (c.f., Lindsay, 2015). However, there is one important difference between the used car market and the scientific publishing market: in the used car market, sellers hold all the cards. Scientific publishing is a community effort, and other stakeholders, such as peer reviewers, can use their influence to effect change.

Imagine if used car dealers were required to have mechanics inspect each car, but they were not required to give the mechanics the information necessary to evaluate the cars, and the mechanics were not paid. The mechanics would rebel. In the scientific market, peer reviewers are the mechanics, and the rebellion is starting (see the Peer Reviewer’s Openness Initiative, opennessinitiative.org; Morey et al., 2016). Because journals cannot function without peer reviewers, and because reviewers are not compensated for their work, they have a great deal of leverage and little to lose by demanding that the system change. The peer reviewers’ demands are quite reasonable. They are not asking for compensation, or indeed for anything that is in their personal self-interest. They are simply asking for transparency, which is necessary for them to do their jobs and evaluate manuscripts effectively.

Some scientists find this revolution in the name of increased transparency and openness distasteful – they do not see a problem with the current system, and fear that this movement will undermine the public’s trust in science. I would argue that these scientists have lost touch with what the public expects of science. For many non-scientists, learning that transparency is not the norm in science comes as a surprise. To anyone outside of the power hubs of science, it must seem obvious that scientists should be held to a higher standard than used car salespeople.

Competing Interests

Simine Vazire is a senior editor at Collabra: Psychology.

Note

¹ Or not. I have no idea what a carburetor is.

References

- Akerlof, G. A.** (1970). The market for “lemons”: Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*, 488–500. DOI: <https://doi.org/10.2307/1879431>
- Bakker, M., van Dijk, A., & Wicherts, J. M.** (2012). The rules of the game called psychological science. *Perspectives on Psychological Science*, *7*, 543–554. DOI: <https://doi.org/10.1177/1745691612459060>
- Brown, N. J. L., & Heathers, J. A. J.** (2016). The GRIM test: A simple technique detects numerous anomalies in the reporting of results in psychology. *PeerJ Preprints*, *4*, e2064v1. DOI: <https://doi.org/10.1177/1948550616673876>
- Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R.** (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews: Neuroscience*, *14*, 1–12. DOI: <https://doi.org/10.1038/nnrn3502>
- Edwards, M. A., & Roy, S.** (in press). Academic research in the 21st century: Maintaining scientific integrity in a climate of perverse incentives and hypercompetition. *Environmental Engineering Science*. DOI: <https://doi.org/10.1089/ees.2016.0223>
- Epskamp, S., & Nuijten, M. B.** (2015). *statcheck*: Extract statistics from articles and recompute p-values. Retrieved from: <http://CRAN.R-project.org/package=statcheck> (R package version 1.0.1).
- Fraley, R. C., & Vazire, S.** (2014). The N-Pact Factor: Evaluating the quality of empirical journals with respect to sample size and statistical power. *PLoS One*, *9*, e109119. DOI: <https://doi.org/10.1371/journal.pone.0109019>
- Freese, J.** (2007). Replication standards for quantitative social science: Why not sociology? *Sociological Methods & Research*, *36*, 153–172. DOI: <https://doi.org/10.1177/0049124107306659>
- Giner-Sorolla, R.** (2012). Science or art? How aesthetic standards grease the way through the publication bottleneck but undermine science. *Perspectives on Psychological Science*, *7*, 562–571. DOI: <https://doi.org/10.1177/1745691612457576>
- Ioannidis, J. P.** (2005). Why most published research findings are false. *PLoS Medicine*, *2*, e124. DOI: <https://doi.org/10.1371/journal.pmed.0020124>
- Kidwell, M. C., Lazarević, L. B., Baranski, E., Hardwicke, T. E., Piechowski, S., Falkenberg, L. S., et al.** (2016). Badges to acknowledge open practices: a simple, low-cost, effective method for increasing transparency. *PLoS Biol.*, *14*(5): e1002456. DOI: <https://doi.org/10.1371/journal.pbio.1002456>
- Lewandowsky, S., & Bishop, D.** (2016). Don't let transparency damage science. *Nature*, *529*, 459–461. DOI: <https://doi.org/10.1038/529459a>
- Lindsay, D. S.** (2015). Replication in Psychological Science. *Psychological Science*, *26*, 1827–1832. DOI: <https://doi.org/10.1177/0956797615616374>
- Lupia, A., & Elman, C.** (2014). Openness in political science: Data access and research transparency. *PS: Political Science & Politics*, *47*, 19–42. DOI: <https://doi.org/10.1017/s1049096513001716>
- Morey, R. D., Chambers, C. D., Etchells, P. J., Harris, C. R., Hoekstra, R., Lakens, D., . . . , Zwaan, R. A.** (2016). The peer reviewers' openness initiative: Incentivizing open research practices through peer review. *Royal Society Open Science*, *3*, 150547. DOI: <https://doi.org/10.1098/rsos.150547>
- Nosek, B., Alter, G., Banks, G., Borsboom, D., Bowman, S., Breckler, S., et al.** (2015). Promoting an open research culture. *Science*, *348*, 1422–1425. DOI: <https://doi.org/10.1126/science.aab2374>
- Piowar, H. A., Day, R. S., & Fridsma, D. B.** (2007). Sharing detailed research data is associated with increased citation rate. *PLoS ONE*, *2*, e308. DOI: <https://doi.org/10.1371/journal.pone.0000308>
- Reidpath, D. D., & Allotey, P. A.** (2001). Data sharing in medical research: an empirical investigation. *Bioethics*, *15*, 125–134. DOI: <https://doi.org/10.1111/1467-8519.00220>
- Simmons, J. P., Nelson, L. D., & Simonsohn, U.** (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*, 1359–1366. DOI: <https://doi.org/10.1177/0956797611417632>
- Simonsohn, U., Nelson, L. D., & Simmons, J. P.** (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*, 534–547. DOI: <https://doi.org/10.1037/a0033242>
- Smaldino, P. E., & McElreath, R.** (2016). The natural selection of bad science. eprint arXiv:1605.09511. DOI: <https://doi.org/10.1098/rsos.160384>
- Tullett, A. M.** (2015). In search of true things worth knowing: Considerations for a new article prototype. *Social and Personality Psychology Compass*, *9*, 188–201. DOI: <https://doi.org/10.1111/spc3.12166>
- van Assen, M. A., van Aerts, R. C., & Wicherts J. M.** (2015). Meta-analysis using effect size distributions of only statistically significant studies. *Psychological Methods*, *20*, 293–309. DOI: <https://doi.org/10.1037/met0000025>
- Vanpaemel, W., Vermorgen, M., Deriemaeker, L., & Storms, G.** (2015). Are we wasting a good crisis? The availability of psychological research data after the storm. *Collabra*, *1*(1), Art. 3, 1–15. DOI: <https://doi.org/10.1525/collabra.13>
- Wicherts, J. M., Borsboom, D., Kats, J., & Molenaar, D.** (2006). The poor availability of psychological research data for reanalysis. *American Psychologist*, *61*, 726–728. DOI: <https://doi.org/10.1037/0003-066X.61.7.726>

Peer Review Comments

The author(s) of this paper chose the Open Review option, and Streamlined Review option, and all peer review comments are available at: <https://doi.org/10.1525/collabra.74.pr>

How to cite this article: Vazire, S. (2017). Quality Uncertainty Erodes Trust in Science. *Collabra: Psychology*, 3(1): 1, DOI: <https://doi.org/10.1525/collabra.74>

Submitted: 26 October 2016 **Accepted:** 18 November 2016 **Published:** 28 February 2017

Copyright: © 2017 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.