

UC Davis

UC Davis Previously Published Works

Title

Taking steps to increase the trustworthiness of scientific research

Permalink

<https://escholarship.org/uc/item/4ch0149m>

Journal

The FASEB Journal, 28(9)

ISSN

0892-6638

Author

Yarborough, Mark

Publication Date

2014-09-01

DOI

10.1096/fj.13-246603

Peer reviewed

Title: “Taking steps to increase the trustworthiness of scientific research”

Author: Mark Yarborough, PhD
Bioethics Program
University of California Davis
4150 V Street, Suite G100
Sacramento CA 95817
916.734.7158 (phone)
916.734.1531 (fax)
mark.yarborough@ucdmc.ucdavis.edu

Running foot: Trustworthy research

Abbreviations

NIH – National Institutes of Health

CTSA – Clinical and Translational Science Award

IRB – Institutional Review Board

IACUC - Institutional Animal Care and Use Committee

FMEA - Failure Modes and Effect Analysis

RCA - Root Cause Analysis

VA - US Department of Veterans Affairs

Abstract

To enjoy the public's trust, the research community must first be clear about what it is expected to do and then avoid the incidents that prevent it from meeting those expectations. Among other things, there are expectations that published scientific results will be reliable, that research has the potential to contribute to the common good, and that research will be conducted ethically. Consequently, the scientific community needs to avoid lapses that prevent it from meeting these three expectations. This requires a strong commitment to trustworthy research practices, as well as mechanisms that diminish lapses that inevitably occur in complex endeavors such as scientific research. The author presents a model to assess the strength of commitment to trustworthy research and explores proven quality assurance mechanisms that can diminish lapses in research injurious to the public's trust. Some mechanisms identify in advance ways that things can go wrong so that steps can be taken to prevent them from going wrong in the first place. Other mechanisms investigate past errors or near misses to discover their causes so that they can be addressed to avoid similar future instances. The author explains why such methods are useful to efforts to promote research worthy of the public's trust.

Introduction

The future of scientific research depends in very large part upon the public's trust. Accordingly, many strategies are being discussed and employed around the world to promote that trust, ranging from more ethics education for researchers to more education for the public about science.(1-4) A prominent effort to build trust here in the United States is the community engagement emphasis that the National Institutes of Health (NIH) required as an initial central tenet of its ambitious Clinical Translational Science Awards (CTSA) program.(5) A recent Institute of Medicine report about the CTSA's heartily endorsed that emphasis.(6) As important as such initiatives are, they are insufficient on their own to sustain the public's abiding trust. As one commentator on trust has noted, the key to enjoying trust is deserving trust.(7) If those of us in the research community who care about the public's trust in our work accept this premise, it will not suffice to limit our efforts to such matters as community engagement or research ethics education. We also need to be sure about how, exactly, we deserve the trust that we continuously solicit. We would all benefit from greater consensus on what constitutes trustworthy research in the first place, as well as what the practices are that help to assure it. Neither of these topics receives much notice, despite the growing attention and fret about the public's trust in research, fret reflected by leading voices in the research community.(8-10) What follows is an attempt to move beyond fret to action, by focusing on two things the research community must do to demonstrate its trustworthiness. First, we need to be clear about what, exactly, our research is expected and trusted to do and, second, we need to employ strategies that avoid or minimize the incidents that prevent it from meeting those expectations.

Being clear about what trustworthy scientific research is

There are at least three things scientific research is trusted to do. First, readers of scientific research expect published results to be reliable. Since the purpose of science is discovery and extension of knowledge about the world, efforts at discovery must be properly conducted and accurately reported so that others can incorporate published discoveries into their own scientific investigations. Unreliable published results, no matter whether they stem from honest error, carelessness, or misconduct, impede this process. Second, very public promises made by modern science generate expectations that research is socially valuable. For example, the American Association for the Advancement of Science states that its mission is to "advance science, engineering, and innovation throughout the world for the benefit of all people."(11) Such statements imply a rational connection between scientific investigations and some aspect of ecological, physical, psychological, or social well-being. Otherwise, research cannot do the good it promises to do. Third, there are ethical and legal

norms that frame how research is supposed to be conducted. Research can be inherently risky and exploitative. As a result, practices meant to protect against ecological, biological, psychological, social, and other risks have evolved. Responsibilities to demonstrate respect for the interests and welfare of the humans and other animals used in research have also been established. These three characteristics of trustworthy research merit a fuller explication elsewhere but this brief discussion of them succinctly shows for our present purposes why research that fails to fulfill any of these basic expectations of *reliability*, *social value*, and *ethical conduct* is research that diminishes the public's trust in science. It also alerts the scientific community to the need to place a premium on avoiding lapses that prevent it from meeting any of these expectations. The remainder of this essay will look at ways to avoid lapses that undermine both the reliability and ethical conduct of research. How to foster greater accountability for the social value of at least some research has been discussed elsewhere and the recommendations found there can be extrapolated to other types of research.(12)

Becoming trustworthy by avoiding incidents that diminish trust

Conversations about trustworthy science need to start with the assurance that exploring concerns about the trustworthiness of research are not meant to call into question the integrity and trustworthiness of the individual people who do research. Instead, the urgency about trustworthy research stems from the complexity of research itself. For example, research now routinely employs advanced technologies and instrumentation, as well as large, multi-site interdisciplinary teams, all of which create ways that things can go wrong. As a result, today's research is prone to multiple shortcomings that have nothing at all to do with misconduct. The traditional safeguards used by the research community – safeguards such as peer review and the protections afforded by the scientific method itself coupled with regulatory compliance - are not able on their own anymore to fend off the many ways that things can and do go wrong in research today. To appreciate the significance of this, consider the extent to which research that relies on extremely large and/or unique data sets, that employs proprietary software or other research tools, or that reports on observations of rare if not unique events renders such traditional safeguards inadequate.(13)

This suggests that new accountability measures are needed to supplement traditional safeguards if the research community is to avoid, or at least diminish, the incidents that undermine trust. One option for identifying such measures is to look to other sectors that also require the public's trust. These sectors implement stringent accountability practices, often in excess of those required by government and other official regulations, to prevent lapses that undermine trust.(14) They rely upon openness and transparency to continually monitor risks and investigate errors in pursuit of safety, quality, and improvement, all of which support trust.

These proactive sectors stand in stark contrast to other domains where people scramble after the fact to redress the harms and reverse the diminution in trust that results from lapses.(15) Too often, this characterizes the approach of the research community. When it focuses on problems that diminish trust, it typically employs ineffective and counterproductive accountability practices characterized by an “excessive reliance on professional expertise.”(14) The result of this approach is a “punitive environment whose focus on individual responsibility leads to secrecy” in much of the research community.(14) The ensuing silence weakens accountability because it conceals rather than prevents potential problems. As a result, trustworthiness is diminished. In contrast, rather than focus on individuals and what they did or did not do, more proactive sectors use transparent systems-based approaches to ask how an incident might or did occur so that it can be avoided in the future.

This brief overview illustrates the benefits the research community could enjoy if it changed its approach to accountability. But change will never occur in the absence of a strong commitment to trustworthiness and the adoption of mechanisms that can promote it. Figure 1 is designed to help us appreciate both why we need to focus more on trustworthiness and how we can achieve it. It is a modification of a previously adapted model(14) to measure the evolution of safety promotion.(15) The new model replaces developmental stages promoting safety with developmental stages promoting trustworthiness. At the lowest point of development is a stage characterized as “Pathological,” indicating a maximum amount of reactivity toward trustworthiness on the part of those in charge. At this stage, one only worries about whether work is trustworthy when something bad happens that causes others to question trustworthiness. In contrast, the highest point of development is a stage characterized as “Generative,” reflecting a collective commitment and ability to conduct work in a way that “generates” trustworthiness.

The model is offered for use as a simple diagnostic that any number of individuals in positions of leadership in scientific research can use. Research team leaders, lab directors, and Principal Investigators could use it. Institutional Review Board (IRB) and Institutional Animal Care and Use Committee (IACUC) directors and other administrative leaders, including Deans and compliance officials, could also use it in their research oversight roles. The model assesses whether or not there are proactive accountability practices in place that both drive and assure a continuous focus on trustworthiness. If there are, it means one is either approaching or at the “Proactive” or “Generative” stages of trustworthiness. If there are not, it reveals that better accountability measures are needed.

At what stage along the path to trustworthiness is scientific research? While this is a question that may produce different answers for different areas of science and which is best directed at specific research teams and

their institutional settings, in general terms one would conclude that much, possibly even most, research has yet to consistently achieve across the board the higher stages of development most conducive to trustworthiness.(14) One reason for this is that the common accountability measures used in so much research fail to place the research community on a sufficiently proactive footing. To illustrate why this is so, let us look briefly at one representative example from biomedical research, the IRB review process. This process involves extensive regulatory requirements meant to protect the safety, welfare, and rights of people who participate in human subjects research. It makes sense to look here because IRBs anchor the regulatory regime pertaining to human subjects research and thus employ the major accountability practices in place to help assure the public that it can trust that human subjects research will be both safe and ethical.

IRB review and oversight focuses extensively on compliance with official regulations and guidance and mainly utilizes review of records and applications to assure compliance. IRBs, for example, spend much time reviewing the content and wording of consent forms, despite the wealth of evidence regarding the ineffectiveness of written consent forms to promote genuine informed consent. For instance, both the length and readability of consent forms have been found to be problematic(16)^(see multiple references at the citation) and the problems research participants have distinguishing between what is research and what is individualized medical care remain entrenched.(17, 18) Given the central role that informed consent plays in the ethical conduct, and thus the trustworthiness, of medical research, a research community at the Proactive or Generative phases of trustworthiness would systematically do more than comply with regulatory requirements.

To be sure, there are many IRBs and many individual research units that go beyond regulatory requirements and work proactively to improve the consent process, but that belies the problem. These additional proactive measures meant to improve quality and thus demonstrate trustworthiness are episodic; they are not uniform. Thus it is chance, not design, which determines when research is “research [we know] we can trust to be ethical.”

This brief example is not meant to negate the importance of either IRB review or the other extensive compliance measures also in place in science to assure both that research is conducted ethically and that published results are reliable. Responsible Conduct of Research education, mandatory disclosures about financial conflicts of interest, and investigations of individuals who engage in misconduct all surely contribute to efforts to assure that research is trustworthy. As important as these efforts are, though, one has to consider whether they are any more capable than IRB review of transitioning the research community to the “Proactive” and “Generative” stages that can continually assure the public that research will be conducted in a manner worthy of its trust. So long as their principal focus is on assigning responsibility to individuals rather than

understanding the causes of events and how to prevent them, and on educating individuals rather than improving systems, then they will fall short of the mark. None of these major accountability practices are capable of routinely generating insights about why things go wrong at times in research. So they are ineffective at preventing and reducing future lapses that can diminish the public's trust in research. Thus, something more, as illustrated below, is needed to improve accountability. Otherwise, no sustained progress toward the "Proactive" and "Generative" stages on the path to trustworthy research can occur.

Taking concrete steps to avoid research that diminishes the public's trust

Progress must start with aspirations to reach the "Proactive" and "Generative" practices that can do so much to promote trustworthy research. Yet as necessary as these high aspirations are, we need more than aspiration alone. Equally important are practical measures that can move us "up" along the path to trustworthy research so that our high aspirations can be realized. That is where we now turn.

When things go wrong, as they inevitably do in any sustained complex activity like scientific research, sectors at the "Proactive" and "Generative" stages of trustworthiness employ quality assurance measures that discover their vulnerabilities to error and mistakes so that they can reduce or eliminate them. Measures like this are needed in the research setting if "science you can trust" is to be more than an aspiration. Auditing is one measure used in multiple settings to improve quality that has been recommended for use in the research setting.(19) There are at least two additional candidate mechanisms that may also be able to help assure both the reliability of scientific publications and the ethical conduct of research. They are Root Cause Analysis (RCA) and Failure Modes and Effect Analysis (FMEA). FMEA is a well-established process that identifies in advance ways that things can go wrong so that steps can be taken to avoid them going wrong in the first place. RCA, on the other hand, is a widely established process used to investigate past errors or near misses to discover their causes so that those can be addressed to avoid future instances. I now briefly explain both, beginning with RCA.

RCA "has been widely adopted as a central method to learn from mistakes."(20) It has been used with great success in such diverse activities as aviation and clinical medicine to reduce error and increase safety. Its success in such different settings is a result of it being a straightforward approach to answer three basic questions: 1) what happened, 2) why, and 3) can anything be done to prevent it from occurring again in the future?(21) If one is interested in preventing problems, it is not enough to understand the "who, what and how" something happened, which are the typical questions asked. Understanding "why it happened... is the key to preventing similar recurrences" and that is the point of RCA.(22)

For example, it is not enough to know that “Maria got paid half of what she was owed on her contract,” the *what*, because “Jim made a mistake in his paperwork,” the *who* and *how*. To prevent the same mistake from occurring again one must also know that Jim made the mistake because the procedures manual he was instructed to use was out of date, the *why*. Knowing that the use of an outdated manual, rather than Jim, is the “root cause” leads to efforts to collect outdated manuals and replace them with updated ones, thereby significantly reducing the chances of future under-payments not only by Jim but by everyone else who processes reimbursements to each and every contract worker. This simple example indicates the key features of RCA. It begins with a gathering of facts, followed by identification of the causes that contribute to the facts under review and ultimately the discovery of the root causes, followed by a set of recommendations for implementation to prevent future recurrences.(22)

Like RCA, FMEA has been used with great success in diverse settings to reduce error and increase quality. In contrast to RCA, though, a FMEA process is used to anticipate ahead of time what problems could happen, i.e., failure modes, so that actions can be taken to prevent them. Not only does it identify the various ways a process might fail, it also determines the severity of the consequences, i.e., effect analysis, of each possible failure so that people can identify and prioritize corrective measures.(23) For example, if Jim’s company had conducted an FMEA on its bulk purchasing practices, it would have discovered that some key procedures and policies are changed more frequently than bulk printed materials are depleted. Thus, it would know that at some point in time the procedures manual would become outdated, resulting in, among other things, inappropriate reimbursement of contract employees. Corrective measures could then be implemented. For instance, the company might switch to electronic publishing of the procedures manual to make sure that manuals are updated more quickly. Or, it could require electronic rather than paper forms for requesting checks so that completed forms would comply with procedures. Both would reduce the chances that people like Maria would be underpaid.

Since FMEA and RCA might help the research community transition to the “Proactive” and “Generative” stages that foster trustworthy research, both warrant investigation to see if they can provide understanding about the myriad ways things do and can go wrong throughout the life of research projects, from the design and conduct of research to the publication of results, and recommend corrective actions that will reduce the chances that lapses occur. For example, research teams and/or the institutional officials where they work could use FMEA to improve team science. They could use both existing research practices and hypothetical cases to troubleshoot how and where problems performing complicated multi-step tasks that need to be completed at multiple sites could arise and then take steps in advance to prevent their occurrence.

Similarly, they could use RCA on actual errors in data entry to identify the causes of them and implement changes based on these findings in an attempt to prevent future similar occurrences.

There are several reasons that suggest that people in the research community would be open to using both FMEA and RCA to progress to the proactive and generative stages of trustworthiness. First of all, researchers are used to a culture of research oversight and review. Whether it is the requirement for biological hazards safety review, IACUC or IRB review, scientific merit review, or even review of alleged misconduct, there are sponsor and institutional practices in place to enforce various standards. FMEA and RCA could join these practices. Similarly, many of the country's biomedical research institutions are already benefitting from the ability of FMEA and RCA to change culture to improve quality and reduce errors in patient care. Hospitals accredited by the Joint Commission, as well as the US Department of Veterans Affairs (VA), already use FMEA and RCA to investigate clinical errors in order to improve quality and safety. Therefore, many research institutions are already familiar with and accept these processes. Expanding FMEA and RCA to the research setting would be a reasonable evolution of an institution's focus on quality improvement. Third, due to mandatory review and reporting requirements, IRBs and IACUCs are required to investigate incidents of suspected non-compliance with regulations and higher education institutions are required to investigate research misconduct. Consequently, research institutions already know where they could target both of these preventive strategies to good effect.

Finally, FMEA and RCA bring a distinct advantage to the table compared to existing approaches to research lapses. The individuals targeted by the current practices to investigate research misconduct allegations frequently view the process as punitive.(24) Since investigations of non-compliance with human subjects regulations similarly target individual investigators, the same punitive, stigmatizing effect likely occurs. FMEA and RCA should avoid much of this negativity because they focus on events, why they occur, and how to prevent them rather than on individual actors and what they may have done wrong. Unlike the current punitive approaches, FMEA and RCA could better bring the research community together in self-reflection and problem-focused activities since they avoid assigning blame and creating stigma. For all these considerations, it makes sense to think that they could make important contributions to the scientific research community.(14, 25)

Challenges

As potentially beneficial as FMEA and RCA might prove to be to the research community, implementing them will be challenging. Since neither would replace mandated investigations of certain types of suspected wrongdoing, institutions must perceive their value to be worth the additional work they will entail.

Individuals may also be reluctant to participate in either. But since both FMEA and RCA focus on understanding the causes of lapses in order to prevent future mishaps, not blaming or exonerating individuals, both should eventually be accepted by individuals, as happened when the VA began to use RCA to address medical errors. There was a noticeable “shift in the root causes identified, blaming individuals less and increasingly attributing the problem to systemic causes.”(20)

The ultimate challenge in the research setting that FMEA and RCA will likely face is getting research teams and institutions to implement the recommendations they will produce. This has been found to be a problem in the clinical arena, where institutions often fail to enact the recommendations stemming from RCAs.(20) The challenge of being willing to implement recommendations reminds us of the significance of aspiring to trustworthiness in the first place.

Concluding Considerations

The more the research community responds after the fact to incidents that diminish trust, the more it leaves to chance the public’s support for its work. Given this reality, it is disconcerting that more effort is not being focused on transformative accountability practices that can eliminate lapses, large and small, that erode the trustworthiness of research. This lack of requisite effort stems at least in part, no doubt, from our tendency in the research community to conflate being trusted with being trustworthy. Thus, we may fail to appreciate that it is possible to enjoy trust without deserving it. The spate of efforts highlighted at the outset to promote the public’s trust in research suggests that we may be placing too much emphasis on being trusted, possibly at the expense of being trustworthy.

Even if these efforts succeed at increasing trust, that increased trust will do nothing to advance the research community itself along the path to the “Generative” stage of trustworthy research. More likely, it will provide both a false sense of security that the trust enjoyed today will also be there tomorrow and a sense of complacency that trustworthy practices themselves are unimportant. The only way to reach the “Generative” stage, where “our work is work we know the public can trust,” is to have standard accountability practices in place, along with a deliberately crafted culture, that assure that level of aspiration and success.

FMEA and RCA, given their proven success in so many other settings, are obvious accountability practices warranting great interest by the scientific community. Their adoption can equip every research team and institution with straightforward mechanisms for realizing a number of benefits. Focus can shift from “problem individuals” to “trustworthy science.” Administrative leaders, compliance officials and researchers

can begin to work together to improve processes and reduce errors. Institutional cultures can become more collaborative. These changes would all bode well for the future of research and the important benefits it pursues.

Acknowledgments: The author would like to thank Jeffrey Elias, PhD, Lee Hilborne, MD, Michael S. Wilkes, MD, MPH, and anonymous reviewers for helpful editorial comments pertaining to portions of the manuscript.

Funding/Support: A portion of Dr. Yarborough's time was supported by the National Institutes of Health, through grant #TR 000002.

References

1. Science and Trust Expert Group Report and Action Plan: Starting a National Conversation about Good Science.
<http://interactive.bis.gov.uk/scienceandsociety/site/trust/files/2010/03/Accessible-BIS-R9201-URN10-699-FAW.pdf>
2. Singapore Statement on Research Integrity.
3. Integrity, E. M. O. F. o. R. Fostering Research Integrity in Europe A report by the ESF Member Organization Forum on Research Integrity.
4. Joint Statement of the II Brazilian Meeting on Research Integrity, Science and Publication Ethics. Vol. 2013
5. Institutional Clinical and Translational Science Award (U54). Vol. 2012
6. Alan I. Leshner, S. F. T., Andrea M. Schultz, and Catharyn T., Liverman, E., Translational, C. t. R. t. C. a., Advancing, S. A. P. a. t. N. C. f., and Medicine, T. S. B. o. H. S. P. I. o. (2013) *The CTSA Program at NIH: Opportunities for Advancing Clinical and Translational Research* The National Academies Press, Washington, D.C.
7. Hardin, R. (2002) *Trust and Trustworthiness*, Russell Sage Foundation, New York
8. (2010) A question of trust. *Nature* **466**, 7
9. Cicerone, R. J. (2010) Ensuring integrity in science. *Science* **327**, 624
10. Basken, P. (2010) NIH Proposes Tougher Rules on Financial Conflicts of Interest. In *The Chronicle of Higher Education* Vol. May 20, 2010
11. Science, A. A. f. t. A. o. Vol. 2010
12. Yarborough, M. (2013) Increasing Enrollment in Drug Trials: The Need for Greater Transparency About the Social Value of Research in Recruitment Efforts. *Mayo Clinic Proceedings* **88**, 442-445
13. Jasny, B. R., Chin, G., Chong, L., and Vignieri, S. (2011) Again, and Again, and Again. *Science* **334**, 1225
14. Yarborough, M., Fryer-Edwards, K., Geller, G., and Sharp, R. R. (2009) Transforming the culture of biomedical research from compliance to trustworthiness: insights from nonmedical sectors. *Acad Med* **84**, 472-477
15. Hudson, P. (2003) Applying the lessons of high risk industries to health care. *Qual Saf Health Care* **12(suppl 1)**, i7-i12.
16. Kass, N. E., Chaisson, L., Taylor, H. A., and Lohse, J. (2011) Length and complexity of US and international HIV consent forms from federal HIV network trials. *Journal of general internal medicine* **26**, 1324-1328
17. Appelbaum, P. S., Roth, L. H., and Lidz, C. (1982) The Therapeutic Misconception: Informed Consent in Psychiatric Research. *Int J Law Psychiatry* **5**, 319-329
18. Kimmelman, J. (2007) The therapeutic misconception at 25: treatment, research, and confusion. *Hastings Cent Rep* **37**, 36-42
19. Shamo, A. E. (2013) Data audit as a way to prevent/contain misconduct. *Accountability in research* **20**, 369-379
20. Wu, A., Lipshutz, A., and Pronovost, P. (2008) Effectiveness and Efficiency of Root Cause Analysis in Medicine. *JAMA* **299**
21. Affairs, U. D. o. V. National Center for Patient Safety: root cause analysis (RCA). Vol. 2007
22. Rooney, J., and Heuvel, L. V. (2004) Root Cause Analysis for Beginners. *Quality Progress* **July**, 45-53

23. Apkon, M., Leonard, J., Probst, L., DeLizio, L., and Vitale, R. (2004) Design of a safer approach to intravenous drug infusions: failure mode effects analysis. *Quality & safety in health care* **13**, 265-271
24. (June 1996) Survey of Accused but Exonerated Individuals in Research Misconduct Cases: final report., Research Triangle Institute, Washington, D.C.
25. Opel, D., Diekema, D., and Marcuse, E. (2011) Assuring research integrity in the wake of Wakefield. *BMJ*, d2