Resisting the Siren Call of Behavioral Change

Author(s)

Kenneth D. Pimple
Douglas Adams

Year

2008

Description

This essay argues against the Federal mandate that requires assessment of behavioral change in responsible conduct of research training programs. True assessment of behavioral change cannot be accomplished by individual instructors, but only on a large scale over times the institutional and national culture of research is gradually changed.

Abstract

The abstract submission form for this conference notes that the organizers "are particularly interested in presentations and panels that will [among other things] … present strategies to assess RCR EIT, with particular attention to assessment of its impact on research behavior."

Given the embarrassment and damage suffered by the research community by widely publicized and outrageous cases of research misconduct, the notion that RCR EIT can change research behavior is admittedly attractive, even seductive. But what kind of assessment is envisioned, and what is meant by RCR EIT's "impact on

research behavior?" We take the phrase "impact on research behavior" to mean change in research behavior, ideally from less-responsible to more-responsible, on the part of recipients of RCR EIT. We outline three broad strategies for assessing behavioral change – a individual assessment (has this particular person's behavior changed?), institutional assessment (has research behavior at our university changed?), and national assessment (is the overall rate of misconduct decreasing?) – and describe strengths and weaknesses of each.

We argue that using behavioral change as a target of individual assessment is likely to be counterproductive because it will invite, if not require, instructors to view their students as irresponsible to begin with, the goal being to redeem them. Viewing one's students as miscreants who need correction is only appropriate when the students have been referred to the educational intervention on the basis of bad behavior.

In particular, we argue against a Federal mandate that makes assessment of behavioral change mandatory for RCR EIT if such a mandate either specifies or turns out to be widely interpreted as requiring individual assessment using a pre-test/post-test model. We argue that such a mandate would be unjustified and would likely have several undesirable, unintended consequences.

Measurements of attitudes and knowledge are more easily assessed and are more appropriate barometers of the success of RCR EIT. Assessment of meaningful change in research behavior cannot be accomplished by individual instructors, but can only be undertaken on a larger scale over time as the institutional and/or national culture of research is gradually changed, with RCR EIT as an important contributor to positive change.

Body

Introduction

In the paper that I presented yesterday, I argued that Federal mandates for education, instruction, and/or training in the responsible conduct of research (RCR EIT) have had undesirable unintended consequences. I shared some ideas about the direction RCR EIT should take, and asked the Federal government to refrain from issuing more mandates in the future. In a similar vein, my co-author Doug Adams and I are about to put forward an argument intended to forestall possible

undesirable unintended consequences of Federally mandated insistence that RCR EIT show behavioral change.

Our concern arises from the statement in the abstract submission form for this conference that the organizers are "particularly interested in presentations and panels that will … [among other things] present strategies to assess RCR EIT, with particular attention to assessment of its impact on research behavior."

We will begin by sharing our understanding of this statement. Then we will provide the reasons it worries us, and conclude with suggestions about how we believe RCR EIT should fit into efforts to promote integrity in research.

Three strategies

We take it that the general nature of the "impact" is clear: It must have to do with whether research behavior becomes more responsible, less responsible, or neither after a course of RCR EIT. The search for the impact of RCR EIT can be guided by at least three different strategies.

First, and perhaps most obviously, the strategic goal might be to demonstrate the impact of a specific intervention on specific individuals. For example, one might offer a two-hour workshop on responsible data management that begins with a pre-test and ends with a post-test designed to show whether the data management behavior of the workshop participants has changed.

We will call this strategy "individual assessment."

A second strategic goal might be to discover the impact of an institution-wide effort to improve research behavior through RCR EIT. For example, a university might start with a comprehensive survey of research behavior, then institute many interventions – workshops, graded courses, seminars, brown-bag lunches, public lectures, and so forth. Some time after the initial survey, probably a few years later, a second survey would be undertaken to assess impact. Here the unit of study is the university rather than the individual researcher, but individual assessment could be one part of the overall design, which we believe would be desirable, but not essential. This is similar to the approach advocated in the Institute of Medicine's report, Integrity in Scientific Research: Creating an Environment that Promotes Responsible Conduct, as a "process of self-assessment and external peer review" to "evaluate and enhance the integrity of [the institution's] research environments,"

ideally coupled with existing accreditation processes (IOM 2002).

We will call this strategy "institutional assessment."

In a third strategy, the impact of RCR EIT might be assessed is on an even larger scale – the scale of the national research enterprise. As the number of researchers who have undergone RCR EIT increases, does the rate of misconduct and questionable research practices fall, stay flat, or rise? Again, this strategy could incorporate both individual assessment and institutional assessment, but we believe it could be accomplished without the use of either.

We will call this strategy "national assessment."

The methods necessary for these three assessment strategies would be very different. Any one of these strategies might be favored on a variety of criteria, and each has its peculiar strengths and poses its own challenges.

Individual assessment

Individual assessment would be the least demanding of the three, but, we believe, the least likely to show behavioral results.

It would be relatively easy because a single instructor could undertake the work on her or his own, while the other two would have to be much larger enterprises. The assessment would be specific to the intervention, which could lead to improvements in the intervention, and it could be disseminated for use by other researchers.

An intervention can only be shown to effect change against some kind of control. When Doug and I think of individual assessment, the first thing to come to mind is comparing an individual after the intervention to that same individual before the intervention on a pre-test/post test model. However, it is also possible, and often desirable, to compare the intervention group to a control group of different individuals who have not had the intervention. (Anderson et al. 2007; Funk, Barrett, and Macrina 2007)

It would be typically impracticable for an individual instructor, or instruction team, to find a control group against which to measure the effectiveness of an RCR EIT intervention, suggesting that the most common way to show behavioral change by using an individual assessment strategy would use a pre-test/post-test model. Positive change in actual research behavior could only be shown if a given student's

initial behavior had already been shown to be irresponsible.

Our knowledge about how many people engage in questionable research practices or commit research misconduct is incomplete,(Martinson, Anderson, and de Vries 2005; Steneck 2002; Swazey, Anderson, and Lewis 1993) but it seems likely that it would not be uncommon for most of the learners taking part in a given RCR EIT intervention to show adequately responsible research behavior in the first place. Even when an instructor is lucky enough to enroll a large number of irresponsible researchers, it will be difficult to effect behavior change with a short intervention such as a Web tutorial or a half-day workshop, and it will be difficult to show behavior change even at the end of a semester-long intervention. It will typically be difficult to get longer-term measures – say 6 months or a year after the end of the course – at all because learners will have moved on and will have little incentive to take part in later assessment efforts.

While most instructors routinely assess student learning, emphasizing behavior change in individual assessment will strike many as impractical. There are much more easily assessed impacts, such as on knowledge and attitudes, that strike us as more appropriate under this strategy.

We argue that behavioral change is most likely to be a counterproductive target of individual assessment, because it will invite, if not require, instructors to view their students as irresponsible to begin with, the goal being to redeem them. Viewing one's students as miscreants who need correction is only appropriate when the students have been referred to the educational intervention on the basis of bad behavior, perhaps by a parole officer. In all other settings it unnecessarily poisons the instructor-student relationship. This also argues for concentrating on assessing individual characteristics that do not carry as much judgmental freight.

Institutional assessment

Institutional assessment would obviously be more demanding than individual assessment. It would require the coordinated cooperation of many individuals and institutional units, including academic departments and research administration offices. It would take longer to design and implement and it would be much more expensive. However, it would also result in many more people receiving RCR EIT and much more data on how their behavior changes, and it would enlist many more people in providing RCR EIT and otherwise committed to the enterprise. If done well,

we think that the effort-to-result ratio would be better than that of individual assessment.

Although data collection on impact would be a much larger endeavor in institutional assessment, it would also more readily allow for a longer interval of time to show change. There would be many more individual components to assess, but we argue that a relatively simple evaluation, much like a course evaluation, would be adequate for most components because the intention would not be to measure the impact of each component, but the general change fostered by all of the components.

From our point of view a major advantage to following an institutional assessment strategy is that RCR EIT would not be the only intervention to be assessed – at least not in any plan that we would consider sensible. We argue that it would be foolish to try to evaluate, change, and re evaluate the research behavior of an entire institution without taking a hard look at all aspects of research behavior, including, for example, the impact of interactions with research administration. In an institution in which research behavior is sub-optimal, it seems likely that some offices, functions, and institutional structures could benefit from reform. As difficult as this might be, it would stand a better chance of making a meaningful and lasting change in research behavior than any number of RCR Web sites, workshops, and courses that exist alongside of, but without meaningful interaction with, an ineffective, corrupt, or corrupting institutional research infrastructure. Furthermore, any effort that proved to be successful could well become a standard for other institutions to emulate or adapt.

National assessment

National assessment would most likely take the form of a large-scale survey of research behavior. We advocate exploring the feasibility of a National Research Misconduct Survey (NRMS), modeled on the highly successful National Crime Victimization Survey. The National Research Misconduct Survey would provide a valid and reliable estimate of the overall prevalence and magnitude of research misconduct as well as the situational elements that potentially facilitate misconduct in the practice of science.

A number of obstacles to implementing the NRMS come readily to mind, such as the complexity of developing an appropriate survey instrument, the expense of

implementing it, variation in participation, and dishonesty on the part of survey respondents. However, similar issues have been explored, addressed, and resolved concerning the measurement of illegal behavior, so we have no reason to believe that these obstacles are insurmountable given adequate commitment by the Federal government to see such a project through. A successful implementation of the NRMS would provide individual and situational data related to research misconduct that could be generalized to the overall research community. This data would show which forms of misconduct are most common and provide a profile of who commits misconduct. This would help us identify occupational roles or settings with the greatest opportunity for misconduct, as well as career phases when researchers are most at-risk to engage in research misconduct. NRMS findings would greatly assist efforts to devise prevention strategies and reform misconduct-prone situational attributes.

Resisting Federally mandated individual assessment

It may be clear by now that we believe that there is most to be gained from an institutional assessment strategy, much to be learned from a national assessment strategy, and much at risk in an individual assessment strategy. We will concentrate on the third for the remainder of this paper.

We should first make it clear that we have no objection to efforts to improve the assessment of RCR EIT as such. We doubt the value of concentrating on behavior change as an outcome measure, but if individual researchers and instructors wish to pursue such a line, we wish them success. However, we strongly object to the Federal government's making assessment of behavioral change mandatory for RCR EIT if such a mandate either specifies or turns out to be widely interpreted as requiring individual assessment using a pre-test/post-test model. We do not know whether the Federal government is inclined in this direction, but if it is, we want to make an argument to resist this impulse now, before it is too late.

We will make two points in our argument. First, we argue that it would be unjustified for the Federal government to require that mandated RCR EIT take a "No Researcher Left Untested" approach. Second, we argue that such a requirement would be unlikely to have a positive outcome and would be very likely to have unintended bad consequences.

Lack of justification

Since interest in RCR EIT initially arose in response to widely-publicized and surprisingly widespread and diverse instances of research misconduct, it is reasonable to suppose that Federal mandates to provide such training were implemented with the goal of preventing bad behavior. It also seems reasonable that the Federal government should refrain from issuing unfunded mandates unless they can be shown that they are effective.Pimple was convinced of this in regard to research ethics education sometime in the early 1990s by Nick Steneck, perhaps accidentally.

We doubt that it is necessary to recount to this audience the long struggle over whether the Federal government has purview over fabrication, falsification, and plagiarism (FFP) and other serious deviations from accepted practice, or FFP alone. The issue was settled in the year 2000 with the publication of the "Federal Policy on Research Misconduct," (OSTP 2000) in which research misconduct is defined as FFP.

This being the case, it seems clear that the Federal government has a strong interest in investigating and preventing FFP, but that its interest in so-called questionable research practices (QRP) is much weaker. If this is correct, showing relevant behavioral change on the individual pre-test/post-test model would require us to show that someone who has already committed misconduct will be prevented from doing so by RCR training. We are arguing that the strong Federal interest in investigating and preventing research misconduct – FFP – justifies Federal mandates to investigate and prevent FFP, but not QRP. Most of components of the PHS Core Instruction Areas for RCR concern QRP, not FFP, and we do not think that a Federal mandate for RCR EIT in those areas is justified.

If we are correct, it would be inappropriate to force anyone who has not committed research misconduct into such a regimen of RCR EIT because it would be impossible to show that his or her behavior had been changed in the desired way. It would be like enrolling a cancer-free person into a clinical trial for cancer treatment.

Likely undesirable consequences

If the Federal government were to decide that it is justified in mandating RCR EIT to prevent FFP and QRP, and that RCR EIT must use individual assessment of the pre-test/post-test model to show a positive impact on research behavior because that is the only way to justify the Federal mandate, we predict several undesirable consequences. We have already alluded to the first, namely that RCR educators

would be forced by the logic of the situation to show empirically that some portion of their students start out as irresponsible researchers.

We have already mentioned the pernicious effect this would have on teacher-student relationships, but we foresee even worse consequences. We have no doubt that instruments can be developed to show behavioral change, no matter the true initial state of behavior in any given cohort of students and no matter the actual behavioral result of the educational intervention. If legitimate-looking instruments cannot be devised, data can be faked, and will be if the stakes are high enough. This is teaching to the test at its worst.

Fulfilling such a "No Researcher Left Untested"-style mandate would be prohibitively expensive in every kind of currency, including money, talent, instructional time, and good will.

Of course it might be worth the price if it were to work, but findings from criminology suggest that it will not, or at least that it will not be the most cost-effective approach. Educational programs that target broad audiences with the intention of preventing criminal behavior, such as the DARE drug-use prevention program, are notoriously ineffective. The cost-benefit ratio is much more favorable when educational or other interventions are targeted at individuals who are known to be at risk. The cost per "student" is higher, but the payoff is disproportionately greater – there's more bang for the buck. To identify at-risk researchers and research environments, something like the National Research Misconduct Survey we mentioned earlier is a necessary precursor.

Conclusion

We are not arguing that RCR EIT as such should not be offered broadly; on the contrary, we believe it should be pervasive. However, we argue that it should not be driven by Federal mandates, and it should not become hostage to a quixotic effort to demonstrate behavioral change in researchers whose behavior is not in question.

For any particular RCR EIT intervention, measurements of attitudes and knowledge are more easily assessed than behavioral change and are more appropriate barometers of success. Assessment of behavioral change on the institutional and national level, coupled with a broad effort to study and reform research situations, are worthy undertakings that should be encouraged.

Thank you for your attention.

References

Anderson, Melissa S., et al. 2007. "What do mentoring and training in the responsible conduct of research have to do with scientists' misbehavior? Findings from a national survey of NIH funded scientists." *Academic Medicine* 82.(9): 853-860.

Funk, Carolyn L., Kirsten A. Barrett, and Francis L. Macrina. 2007. "Authorship and publication practices: Evaluation of the effect of responsible conduct of research instruction to postdoctoral trainees." *Accountability in Research* 14: 269-305.

Martinson, Brian C., Melissa S. Anderson, and Raymond de Vries. 2005. "Scientists behaving badly." *Nature*: 737-738. OSTP, (Office of Science and Technology Policy). 2000. "Federal policy on research misconduct." Federal Register 65.(235): 76260-76264.

Steneck, Nicholas H. 2002. "Assessing the integrity of publicly funded research." In Investigating Research Integrity: Proceedings of the First ORI Research Conference on Research Integrity. Eds. Nicholas H. Steneck and Mary D. Scheetz, pp. 1-16. Washington, D.C.: Office of Research Integrity, National Institutes of Health.

Swazey, Judith P., Mellisa S. Anderson, and Karen S. Lewis. 1993. "Ethical problems in academic research." *American Scientist* 81.(6): 542-553.

Notes

A significantly shortened version of this paper was presented at the first biennial conference on "Responsible Conduct of Research (RCR) Education, Instruction, and Training" (http://epi.wustl.edu/epi/rcr2008.htm) St. Louis, Missouri, April 18, 2008, sponsored by the Office of Research Integrity and the Washington University School of Medicine.

Rights

Use of Materials on the OEC

Resource Type

Essay

Parent Collection

Ken Pimple Collection

Topics

Teaching Ethics

Discipline(s)

Teaching Ethics in STEM
Research Ethics