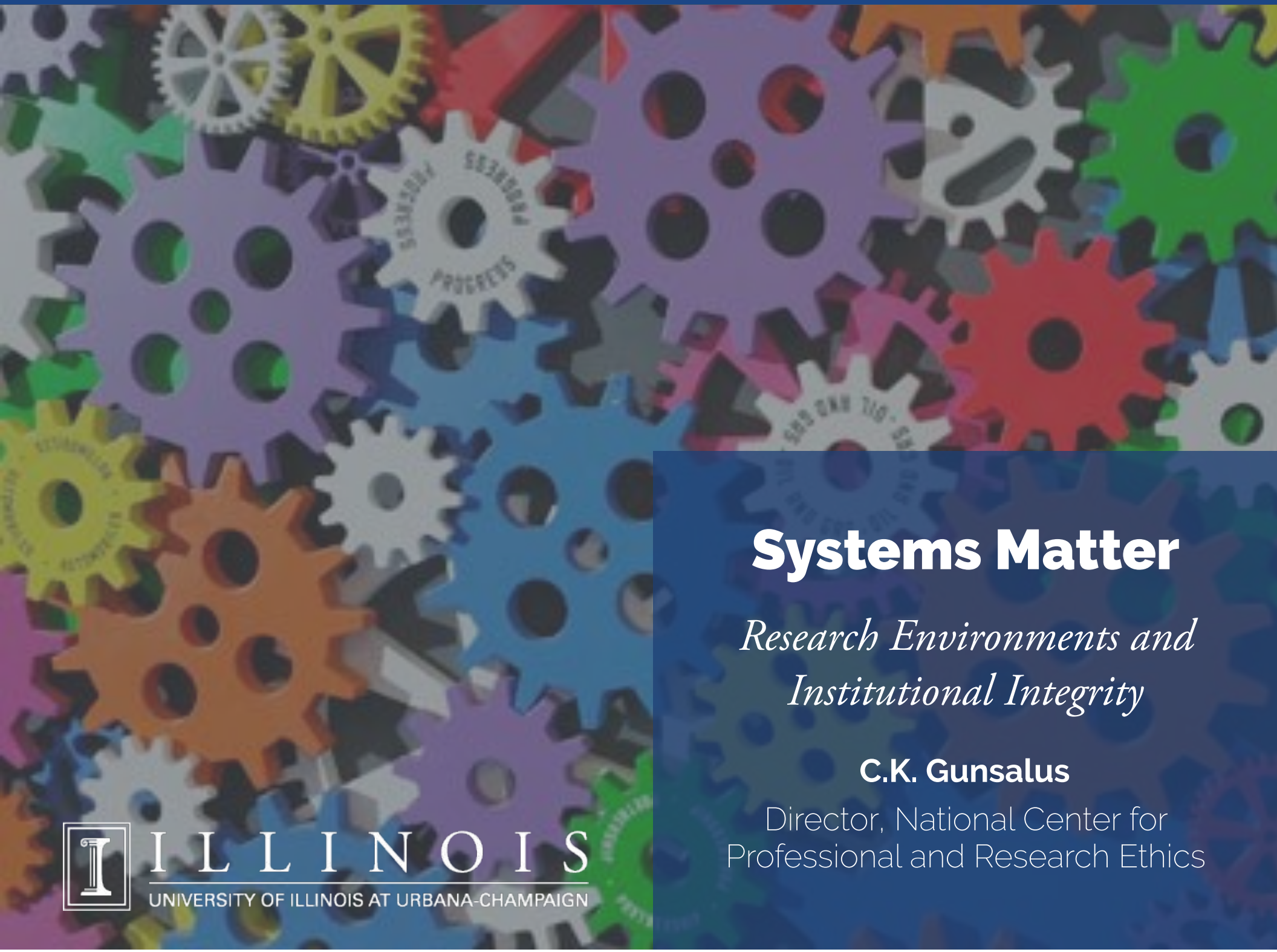


Systems Matter

Research Environments and Institutional Integrity

C.K. Gunsalus

Director, National Center for Professional and Research Ethics

ILLINOIS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

Focus on integrity mindset in research environments to reinforce rigor and reliability.

1

2

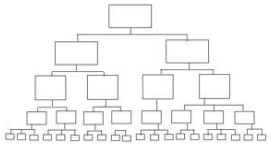
3

Argument

A longer view is instructive

People are complicated

Context matters



Institutional research environments can be a part of the problem—and should be part of the solution

- RIO: research misconduct allegations, investigations
- workplace violence team
- dysfunctional academic units
- degree revocation
- department head, dean training, support
- misuse federal $$
- grievance system oversight discrimination harassment
- bullyproofing academic units
- campus complaint system
- internal investigations
- IRB director, after breach


INTEGRITY
IN SCIENTIFIC RESEARCH
Promoting Integrity in the
Next Generation of Researche
A Curriculum for Responsible Conduct of Research
in Occupational Therapy (2005)
INSTITUTE
OF
MEDICINE
The Responsible Conduct
of Research
in the Health Sciences
AAU
TASK FORCE ON
RESEARCH
ACCOUNTABILITY
Report on
Individual and Institutional Finan
flict of Interest
RESPONSIBLE
SCIENCE
We've been talking about some of the
same elements affecting the rigor and
integrity of research for a long time.
HEARING
SUBCOMMITTEE ON
INVESTIGATIONS AND OVERSIGHT
COMMITTEE ON
SCIENCE, SPACE, AND TECHNOLOGY
U.S. HOUSE OF REPRESENTATIVES
ONE HUNDRED FIRST CONGRESS
FIRST SESSION
JUNE 28, 1989
[No. 73]
Printed for the use of the
Committee on Science, Space, and Technology
Statistical
Challenges in
Assessing and
Fostering the
Reproducibility of
Scientific Results
Integrity and Misconduct
in Research
Report of the Commission on Rese
Integrity
Principles for Protecting Integrity
In the Conduct and Rep
Protecting Patients, Preserving
Integrity, Advancing Health:
Accelerating the Implementation of COI
Policies in Human Subjects Research
2011 Annual Meeting
International Research Con
125th ANNIVERSA
SIGMA X
1886
THE MISSING "R":
REPRODUCIBILITY IN A
CHANGING RESEARCH
LANDSCAPE
A Workshop of the Roundtable on Science
and Welfare in Laboratory Animal Use
(an ILAR Roundtable series)

1986

Science, Statistics, and Deception

John C. Bailar III

Annals of Internal Medicine, 1986, 104, pp. 259–260

Table 1. *Some Practices that Distort Scientific Inferences*

Failure to deal honestly with readers about nonrandom error (bias)
Post hoc hypotheses
Multiple comparisons and data dredging
Inappropriate statistical tests and other statistical procedures
Fragmentation of reports
Low statistical power
Suppressing, trimming, or "adjusting" data; or undisclosed repetition of "unsatisfactory" experiments
Selective reporting of findings

1986

Science, Statistics, and Deception

John C. Bailar III

Annals of Internal Medicine, 1986, 104, pp. 259–260

Table 1. *Some Practices that Distor*

Failure to deal honestly with readers ab
Post hoc hypotheses
Multiple comparisons and data dredgin
Inappropriate statistical tests and other
Fragmentation of reports
Low statistical power
Suppressing, trimming, or "adjusting" d experiments
Selective reporting of findings

1996

An Introduction to Research Ethics*

Paul J. Friedman,
University of California San Diego School of Medicine, USA

Keywords: scientific/research integrity, scientific/research misconduct, research ethics, research fraud, authorship

TABLE 1

Research Activities in which Practical Ethical Problems Arise

Data:
- recording and retaining experimental data
- replication (avoid "cutting corners" or taking shortcuts)
- selecting data for publication or presentation
- analysis, including statistics
- sharing of data and research materials
- ownership of records and ideas
- graduate and postdoctoral student rights

Results:
- statistical analysis not done or reported
- premature use in grants (unconfirmed or best results quoted)
- anticipation of results in abstracts (reported experiments not completed)
- exaggerating significance of results (public or scientific deception)
- self-deception ("mythical thinking") about results or their significance

Publication pressures:
- academic practices that favor long bibliographies
- journal practices that favor short, positive reports
- credit and responsibility should be inseparable

1986

Science, Statistics, and Deception

John C. Bailar III

Annals of Internal Medicine, 1986, 104, pp. 259–260

Table 1. *Some Practices that Distor*

Failure to deal honestly with readers ab
Post hoc hypotheses
Multiple comparisons and data dredgin
Inappropriate statistical tests and other
Fragmentation of reports
Low statistical power
Suppressing, trimming, or "adjusting" d experiments
Selective reporting of findings

1996

An Introduction to Research Ethics*

Paul J. Friedman,
University of California San Diego School of Medicine, USA

Keywords: scientific/research integrity, scientific/research misconduct, research ethics, research fraud, authorship

TABLE 1

Research Activities in which Practical Ethical Problems Arise

Data:
recording and retaining experimental da
replication (avoid "cutting corners" or t
selecting data for publication or present
analysis, including statistics
sharing of data and research materials
ownership of records and ideas
graduate and postdoctoral student rights
Results:
statistical analysis not done or reported
premature use in grants (unconfirmed o
anticipation of results in abstracts (repo
exaggerating significance of results (pu
self-deception ("mythical thinking") ab
Publication pressures:
academic practices that favor long bibli
journal practices that favor short, positi
credit and responsibility should be insep

2013

High-profile studies typically fail at multiple levels:

Begley's six criteria for judging scientific reports:

1) Were studies blinded?
   Almost never
2) Were all results shown?
   Typically not — "representative examples" & data selection bias; western blots that show only a slice; no size markers
3) Were experiments repeated?
   Often not — westerns/immuno-precipitation usually only performed once; typically only use 1/2 siRNAs and in 1/2 cell lines; confusion between replicates and independent experiments
4) Were positive and negative controls shown?
   Typically not
5) Were reagents validated?
   Frequently not — IHC with a polyclonal anti-peptide Ab; small molecule inhibitors
6) Were the statistical tests appropriate?
   Typically not

Nature 497, 433-434, 23 May 2013

1997

EDITORIAL

Ethics: Sending Out the Message

It seems ironic that it took a federal requirement to stimulate formal instruction in ethical conduct in U.S. graduate science training programs.

After all, the ethical conduct of research is central to the integrity of universities, where research and graduate education are inseparable. Yet a requirement for such instruction as a condition of receipt of National Institutes of Health institutional grant funds* may have focused attention on ethical issues more effectively than our own degree programs and more forcefully than a small but steady stream of instances in which researchers violated the norms of professional conduct. These instances range from the spectacular (mouse-painting or outright concoction of data) to the mundane (duplicate publication or guest authorship).

Many researchers feel that the attention devoted to cases of scientific misconduct is disproportionate to their importance and rate of occurrence. But the seriousness of these instances when they do occur, and institutions' often inadequate responses to them, have generated public and congressional skepticism about the ability—and willingness—of the academic community to set and enforce ethical standards. These instances can also endanger morale and foster cynicism among those we are educating.

The relationship between the formal academic curriculum and the informal curriculum that students absorb in hallways, laboratories, and hospitals bears careful examination. What does the informal curriculum teach students in your laboratory or university? What messages do students pick up about authorship and publication practices? How do they see mentors reconcile a desire for a hefty publication record with admonitions not to engage in "salami science" or divide work into "least publishable units"? Do students observe professors maintaining confidentiality in reviewing grant applications? What sort of example do you set? Most important, do the rules apply to everyone in your environment or only to the students?

Somehow it has come to seem unfashionable, almost priggish, to talk about concepts of honor, duty, and obligation. At the same time, increasing funding pressures have created perceived incentives to behave unethically or unprofessionally. What are the boundaries of ethical conduct, and whose job is it to set them and make them stick? Self-policing is a difficult task that few professions seem to have mastered. But the fact that it is difficult doesn't mean that our community shouldn't try. Each of us has the obligation to confront ethical issues and their implications for our personal conduct. Each of us also has the duty to address the ethical aspects of our work with colleagues and students. Institutions have a responsibility to articulate standards for ethical conduct and to see that they are put into practice. These standards should be higher than the merely legal; they should define professional standards of behavior. The goal should be to provide guidance to the well-intentioned—those who may want to do the right thing but who genuinely do not know what is right in a complex situation or how to determine it.

Collaborators should follow the guidelines of leading journals on assignment of authorship. Senior researchers have a special responsibility for seeing that credit is assigned appropriately. Salvador Luria and Max Delbrück set an example well worth emulating in a footnote to their famous 1943 paper.† It read, "Theory by M.D., experiments by S.E.L." Finally, researchers and their universities must accept responsibility for creating an environment in which ethical issues are ordinary topics of conversation and in which ethical conduct is commonplace among leaders and expected of all.

Action from Washington should not be necessary for universities to take these issues seriously or to ensure that their students and trainees learn about ethical conduct as an integral part of their professional education. We must articulate and observe ethical standards in order to avoid an erosion of public trust and the potential for more onerous regulation. It also happens to be the right thing to do.

C. K. Gunsalus

The author is an associate provost at the University of Illinois at Urbana-Champaign. She recently concluded a 4-year term of service as chair of the AAAS Committee on Scientific Freedom and Responsibility.

*"Requirement for Programs on the Responsible Conduct of Research in National Research Service Award Institutional Training Programs," NIH Guide for Grants and Contracts **18**, 45 (1989). A more recent notice appears in ibid. **23**, 23 (1994). †S. Luria and M. Delbrück, Genetics **28**, 491 (1943).

Rethinking Unscientific Attitudes About Scientific Misconduct

By C. K. Gunsalus

MANY AMERICAN SCIENTISTS are fed up with press reports and questions from Congress and the public about scientific misconduct. The concern is drastically overblown, they say, and the government should spend less time and money investigating the few bad apples and concentrate on expanding appropriations for research. After all, some of the most highly publicized charges of misconduct eventually have been dismissed, these scientists note. Relatively few scientists have been found guilty of misconduct, so no elaborate investigative apparatus or intrusive federal rules are needed.

These feelings seem heartfelt and widely shared. What's worrisome is how unscientific they appear.

What's unscientific? Well, it's unscientific to make repeated assertions that scientific misconduct is an extremely small or non-existent problem when we have few or no reliable data supporting those claims. In an extreme example, a 1987 editorial in *Science* said: "99.99999% of all published reports are truthful and accurate, often in rapidly advancing frontiers where accurate data are difficult to collect."

There is no basis for this claim, despite the air of scientific precision conferred by the four digits following the decimal point. Then (as now) we had no direct data on the accuracy of the scientific literature. We simply do not know whether a lot or just a little untruthful information is published. In fact, many scientists vehemently objected a few years ago to a proposed experiment to gather anonymous data on the prevalence of gross misconduct in biomedical research. In the absence of such data, scientists are not exempt from the normal requirement that they be accurate in their public statements.

Moreover, think about the implications of the argument that because scientific misconduct is rare, government does not need regulations and an apparatus to respond. How would the public react to the thesis that because counterfeiting is rare, laws against it and facilities for testing suspect currency cannot be justified?

It's also unscientific to make repeated assertions about the causes of scientific misconduct. Here, too, we lack data. Yet the literature is awash with pronouncements. Typical is a report in *Chemical & Engineering News* of a session at the 1996 meeting of the American Chemical Society in which one panelist asserted: "'But 'fraud in science' is not a real problem. That is because of the psychology of the perpetrators of fraud, and the self-checking nature of the system. The psychopathology of fraud is such that its perpetrators hardly ever contain themselves to manufacturing routine data. Instead, they doctor something important."

What are "routine data"? How does a chemist understand the psychological mindset of perpetrators of fraud without conducting research into the issue? Why are accomplished scientists speaking without evidence to support their assertions? The answer, I believe, is that some structural aspects of universities lead top scientists to minimize the existence of problems and to ignore the possibilities for misconduct that are inherent in research.

The first structural issue is what I call the paradox of the university: A good one is organized so that the active scientists are insulated from what it takes to run it, so that they can think creatively and do science. Productive scientists complain that they are plagued with administrative work and committees, but most of that work is focused on matters directly related to their professional lives—selecting their students and colleagues, and supervising research facilities. Very little is focused on the nitty-gritty of running a large enterprise: what it takes to turn the lights on every day, do the paperwork required by government agencies and foundations, pay the bills, dispose of hazardous wastes, or respond to the odd conduct of troubled individuals.

> "The leaders of science need to be more realistic about the nature of the enterprise that they supervise and defend."

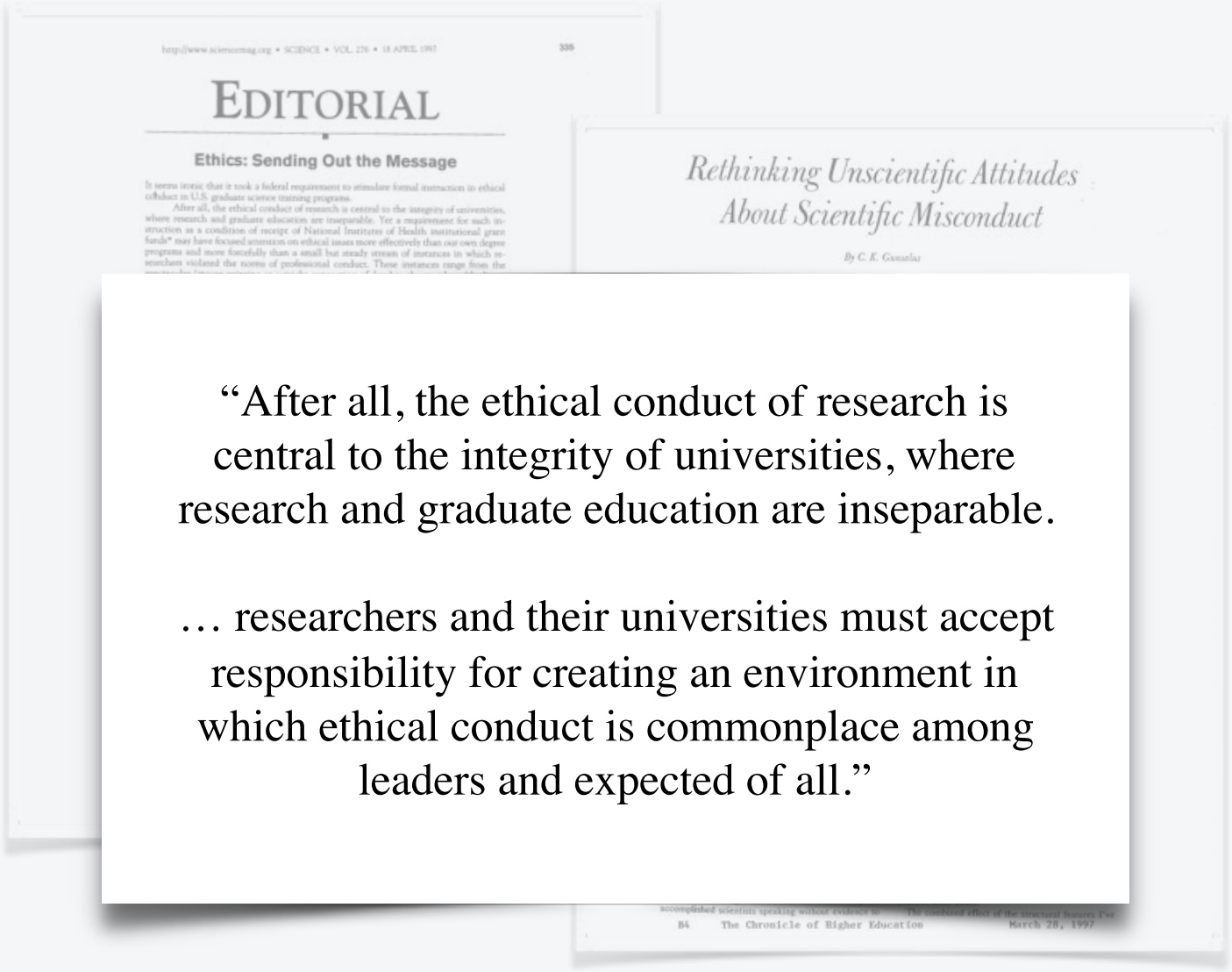
For the most part, this system operates as intended, so that working scientists can, in fact, remain naïve about the realities of day-to-day problems outside their labs. So it's natural that they fail to appreciate the need for rules and systems to deal with those problems. But it doesn't mean those rules and systems aren't necessary.

The second structural issue can be called the bias of the best. In their professional lives, the best people in an institution, particularly the best scientists with exemplary standards of conduct, typically associate only with other top scientists and outstanding students. They normally don't deal much with more-ordinary colleagues, including those whose work ethics or standards may be problematic.

They also have the power, when they do encounter misconduct, to handle problems efficiently. Consider a recent, well-publicized case. When Francis S. Collins, the highly respected director of the National Center for Human Genome Research, found last year that a junior researcher had concocted data, he promptly retracted five published papers on leukemia. The length of the formal procedures to pin down the fraud and respond to it can be measured in months in that case, compared to years in other cases.

The combined effect of the structural features I've

1997



EDITORIAL

Ethics: Sending Out the Message

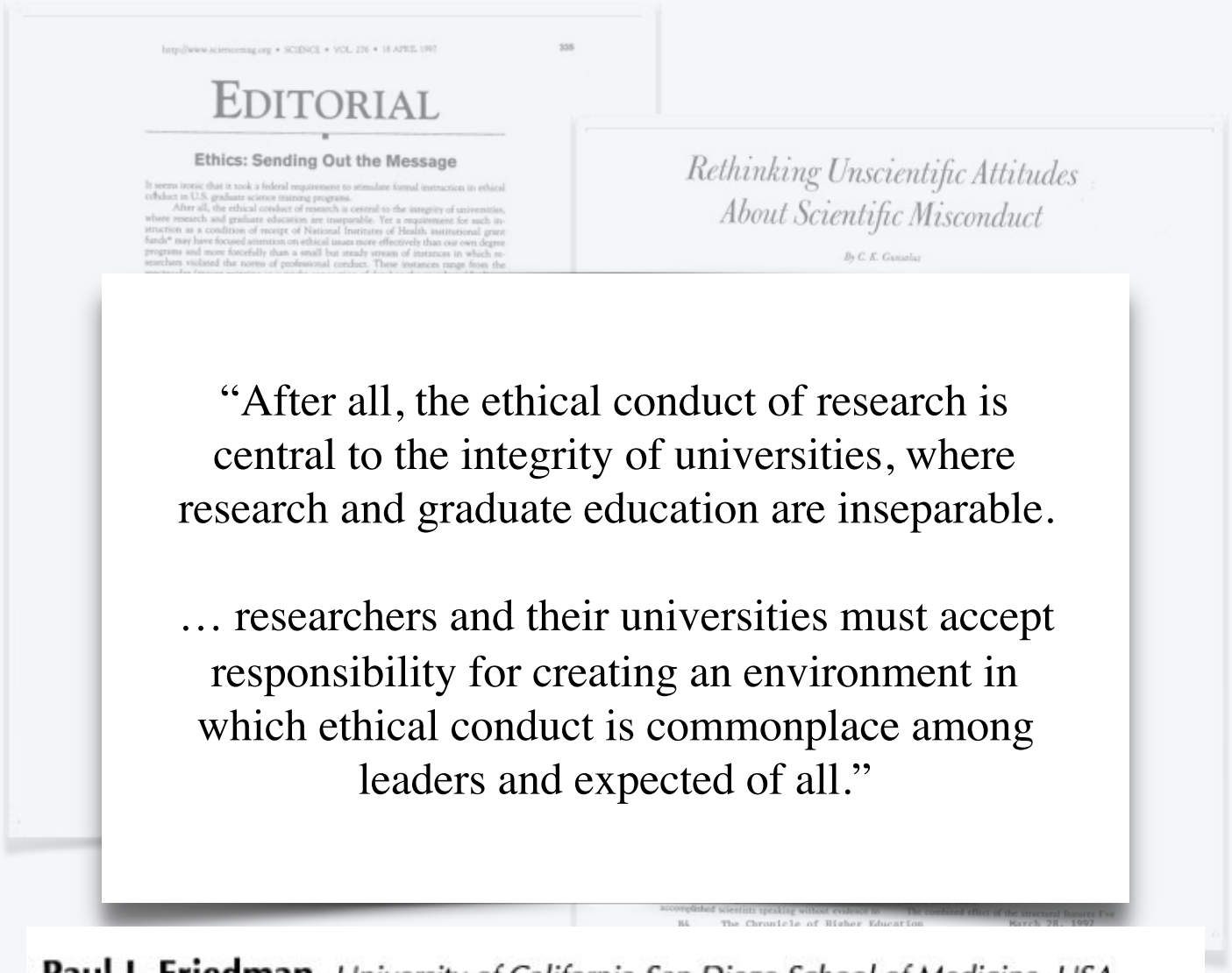
Rethinking Unscientific Attitudes About Scientific Misconduct

By C. K. Gunsalus

"After all, the ethical conduct of research is central to the integrity of universities, where research and graduate education are inseparable.

… researchers and their universities must accept responsibility for creating an environment in which ethical conduct is commonplace among leaders and expected of all."

1997



EDITORIAL

Ethics: Sending Out the Message

Rethinking Unscientific Attitudes About Scientific Misconduct

"After all, the ethical conduct of research is central to the integrity of universities, where research and graduate education are inseparable.

… researchers and their universities must accept responsibility for creating an environment in which ethical conduct is commonplace among leaders and expected of all."

1999

Paul J. Friedman, *University of California-San Diego School of Medicine, USA*

Science and Engineering Ethics (1999) ***5****, 177-178*

One must not ignore another important influence on research integrity: the research environment. Unfortunately, momentum is in the opposite direction

Why is so much work not reproducible?

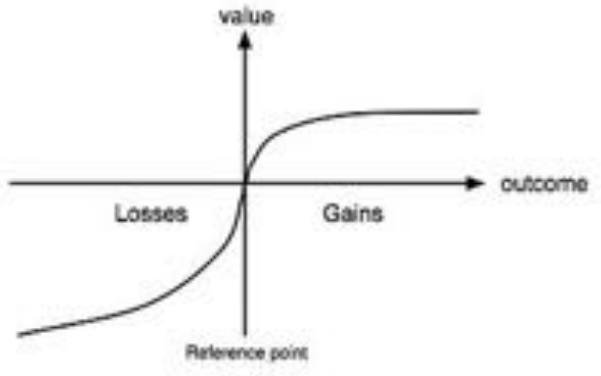
What systems can we put in place to improve reproducibility?


ASK MORE
QUESTIONS

What do we know about cognition and decision-making that contribute to this situation?

How can we use what we have learned from advances in under-standing, and from experience, to reinforce integrity in research environments?

Cognitive Biases



Loss Aversion

Career TRAGEDIES

Temptation
Rationalization
Ambition
Group, Authority Pressure
Entitlement
Deception
Incrementalism
Embarrassment
Stupid Systems

1995

COMMENTARY

by Joshua Lederberg

Sloppy Research Extracts A Greater Toll Than Misconduct

There has been much ado about fraud in science, and even more misunderstanding about its eventual importance in the efficient conduct of science, and our ability to police it. A report on a survey by Judith Swazey was headlined in the *New York Times*"... the myth that fraud in science is a rarity" (L. K. Altman, Nov. 23, 1993, page C3). In fact, as the text of the story took pains to emphasize, the study found that a majority of interviewees had heard of an example—in other words, that some fraud had been found out. We have no idea how often the same examples were in mind. It is small comfort that such surveys in other fields, not excluding politics, law, and journalism, would give equal or larger returns.

The promulgation of fraud is an outrage, striking at the moral roots of the scientific enterprise. But its moral stridency is large, I submit, compared to its practical importance in most scientific fields. A much larger toll is exacted from inadequate experimental design and sloppy execution. The lost effort that is expended in straightening out muddy claims, or merely in plowing through their presentation in the literature, greatly exceeds what can be attributed to intentional fraud.

We do not rely entirely on the intrinsic virtue of the scientific personality. We are all human, and not equally socialized into the deepest respect for the truth, nor equally well trained to avoid even simple logical and statistical fallacies. It is the scientific system of organized skepticism, to borrow Robert K. Merton's phrase, that maintains the integrity of the enterprise.

Scientific claims enter into a cognitive network of great complexity. Rarely is work exactly replicated. Often, that would be extremely difficult to do for practical reasons: localized phenomena, availability of reagents, meticulousness of protocols. But if the work is of any real significance it will be built upon as the basis for further progress, and therein is the ongoing test of the validity of the original findings. When there are discrepancies in the further edifice, that is the most likely occasion for a retest, a close replication. In high-flown, competitive areas, there is some likelihood of duplication of investigation and contest for priority claims. Fraud may then eventuate more on the claims for credit than on the validity of the underlying observations.

There are important exceptions to the aforestated generalizations: when the outcome of scientific investigation is entry not into the network of scientific knowledge, but straight into policy. In clinical investigation, particularly, there are grave public consequences, and the costs of objective replication are prohibitive. So I endorse the clamor that special safeguards be applied with respect to conflict of interest, exposure of raw data, and so forth in that area. Sadly, many clinical investigators are simply not well indoctrinated in the procedural rigor entailed by scientific enquiry.

My own experience over the past 50 years has been that the loss of efficiency in science (for the reasons mentioned) is a hundredfold greater from egregious sloppiness in experimental design, in self-delusion, and in confused reporting than it is from intentional fraud. Few and far between—one might even say celebrated—are the cases in which fraud has really been much of a diversion in scientific progress. In fact, much has been made of historic cases of selective extension of data—Michelson, Mendel—for which there really has been no dissent about the substantiality of the final claims.

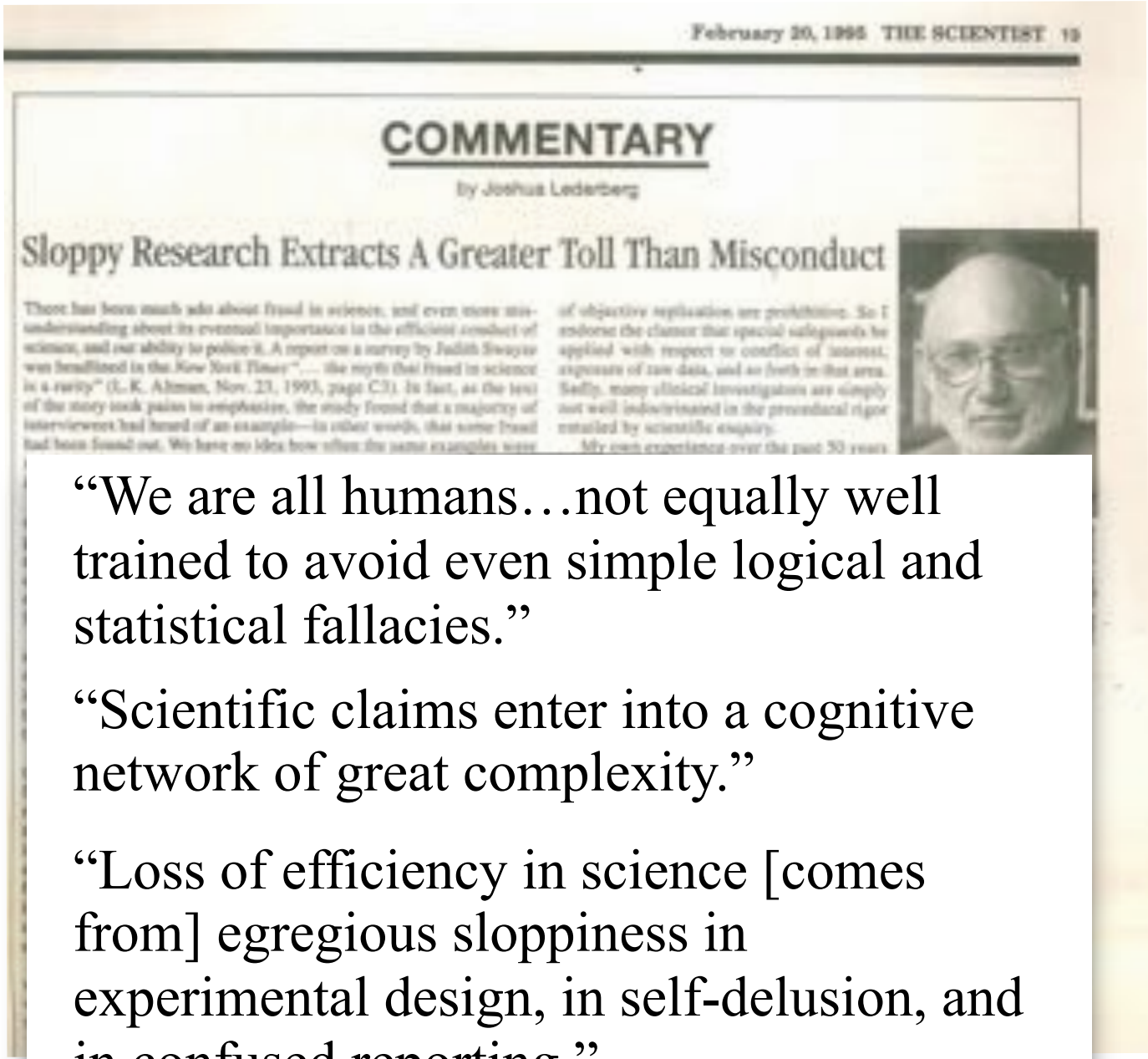
So why don't we put more energy into maximizing the rigor and quality of science as it is practiced? John C. Bailar has commented on efforts to enhance the sophistication of statistical criticism (see his paper "Science, statistics and deception," *Annals of Internal Medicine*, 104:259-60, 1986). A British group led by J.C. Wyatt has designed an expert system, "Design a Trial," to assist in the critique of experimental designs in clinical trials (*Computer Methods and Programs in Biomedicine*, 43:283-92, 1994). I would challenge other critics of science to provide more proof-checking paradigms in the broader conduct of science. And that would help to deal with fraud, as well.

New technological modes of publication may help to answer these faults, especially in assisting the critical discourse that is essential to effective skepticism. There is too high a threshold for publishing a critical note in answer to a perceived error in last month's journal. We are also deterred by the assumption that surely "someone else would have picked it up."

The looming advent of electronic media for such discourse gives an easy technical solution to these barriers to the fulfillment of the Republic of Science.

Joshua Lederberg is University Professor at Rockefeller University.

1995



February 20, 1995 THE SCIENTIST 13

COMMENTARY

by Joshua Lederberg

Sloppy Research Extracts A Greater Toll Than Misconduct

There has been much ado about fraud in science, and even more misunderstanding about its eventual importance in the efficient conduct of science, and our ability to police it. A report on a survey by Judith Swazey was headlined in the *New York Times* "... the myth that fraud in science is a rarity" (L. K. Altman, Nov. 23, 1993, page C3). In fact, as the text of the story took pains to emphasize, the study found that a majority of interviewees had heard of an example—in other words, that some fraud had been found out. We have no idea how often the same examples were

of objective replication are prohibitive. So I endorse the clamor that special safeguards be applied with respect to conflict of interest, exposure of raw data, and so forth in that area. Sadly, many clinical investigators are simply not well indoctrinated in the procedural rigor entailed by scientific enquiry.

My own experience over the past 50 years

"We are all humans…not equally well trained to avoid even simple logical and statistical fallacies."

"Scientific claims enter into a cognitive network of great complexity."

"Loss of efficiency in science [comes from] egregious sloppiness in experimental design, in self-delusion, and in confused reporting."

"[You] are the easiest person to fool. So you have to be very careful about that. After you've not fooled yourself, it's easy not to fool other scientists..."

Richard Feynman, 1974

"Nothing is easier than self deceit."

Demosthenes, 3rd Century, BC


Problems ≠ Bad or careless people

Cell Reports
Commentary

Sorting Out the FACS: A Devil in the Details

William C. Hines,[1,5,*] Ying Su,[2,3,4,5,*] Irene Kuhn,[1] Kornelia Polyak,[2,3,4,5] and Mina J. Bissell[1,5]
[1]Life Sciences Division, Lawrence Berkeley National Laboratory, Mailstop 977R225A, 1 Cyclotron Road, Berkeley, CA 94720, USA
[2]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, MA 02215, USA
[3]Department of Medicine, Brigham and Women's Hospital, Boston, MA 02115, USA
[4]Department of Medicine, Harvard Medical School, Boston, MA 02115, USA
[5]These authors contributed equally to this work
*Correspondence: chines@lbl.gov (W.C.H.), ying_su@dfci.harvard.edu (Y.S.)
http://dx.doi.org/10.1016/j.celrep.2014.02.021

The reproduction of results is the cornerstone of science; yet, at times, reproducing the results of others can be a difficult challenge. Our two laboratories, one on the East and the other on the West Coast of the United States, decided to collaborate on a problem of mutual interest—namely, the heterogeneity of the human

of studying cells close to their context in vivo makes the exercise even more challenging.

Paired with in situ characterizations, FACS has emerged as the technology most suitable for distinguishing diversity among different cell populations in the mammary gland. Flow instruments have

breast reduction mammoplasties. Molecular analysis of separated fractions was to be performed in Boston (K.P.'s laboratory, Dana-Farber Cancer Institute, Harvard Medical School), whereas functional analysis of separated cell populations grown in 3D matrices was to take place in Berkeley (M.J.B.'s laboratory,

Sure, there are **bad** apples

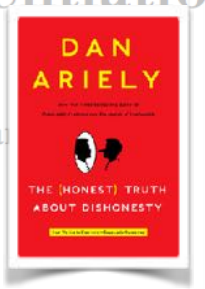
We are each always individually responsible for our own actions.


And, the barrel shapes
perceptions and choices.

Research tells us:

Emphasis on performance	Emphasis on mastery
High stakes	Frequent, low-stakes assessments
Extrinsic motivation	Intrinsic motivation
Low expectation of success	Path to success
Peer culture that accepts	Peer culture that disapproves


Consider the environment.
We know that people are influenced by the choices of those around them.



Research tells us:

"The amount of cheating in which human beings are willing to engage depends on the structure of our daily environment."

The Truth About Dishonesty, Ariely 2013

Academic Environment

Star system

Rewards & incentives

Grey areas in norms

Flawed problem reception and resolution systems

Too Many Environments

Mixed messages

Results, not process

Bad examples

Uneven mentoring

Abuses of power

Problem-solving resources

Suppression of concerns

Retaliation

Social interactions in **LARGE** systems are more complex

PERSPECTIVE

Rescuing US biomedical research from its systemic flaws

Bruce Alberts[a], Marc W. Kirschner[b], Shirley Tilghman[c,1], and Harold Varmus[d]

[a]Department of Biophysics and Biochemistry, University of California, San Francisco, CA 94158; [b]Department of Systems Biology, Harvard Medical School, Boston, MA 02115; [c]Department of Molecular Biology, Princeton University, Princeton, NJ 08540; and [d]National Cancer Institute, Bethesda, MD 20892

Edited by Inder M. Verma, The Salk Institute for Biological Studies, La Jolla, CA, and approved March 18, 2014 (received for review March 7, 2014)

The long-held but erroneous assumption of never-ending rapid growth in biomedical science has created an unsustainable hypercompetitive system that is discouraging even the most outstanding prospective students from entering our profession—and making it difficult for seasoned investigators to produce their best work. This is a recipe for long-term decline, and the problems cannot be solved with simplistic approaches. Instead, it is time to confront the dangers at hand and rethink some fundamental features of the US biomedical research ecosystem.

graduate education | postdoctoral education | federal funding | peer review

By many measures, the biological and medical sciences are in a golden age. That fact, which we celebrate, makes it all the more difficult to acknowledge that the current system contains systemic flaws that are threatening its future. A central flaw is the long-held assumption that the enterprise will constantly expand. As a result, there is now a severe imbalance between the dollars available for research and the still-growing scientific community in the United States. This imbalance has created a hypercompet-

DNA sequencing, sophisticated imaging, structural biology, designer chemistry, and computational biology—has led to impressive advances in medicine and fueled a vibrant pharmaceutical and biotechnology sector.

In the context of such progress, it is remarkable that even the most successful scientists and most promising trainees are increasingly pessimistic about the future of their chosen career. Based on extensive observations and discussions, we believe that these concerns are justified and

doubling of the NIH budget ended, the demands for research dollars grew much faster than the supply. The demands were fueled in large part by incentives for institutional expansion, by the rapid growth of the scientific workforce, and by rising costs of research. Further slowdowns in federal funding, caused by the Great Recession of 2008 and by the budget sequestration that followed in 2013, have significantly exacerbated the problem. (Today, the resources available to the NIH are estimated to be at

Mixed Messages

- Hyper competition
- Irresponsibility rewarded (counting papers, H factors)
- RCR low priority, status, funding
- "Responsible research training" is too compliance-focused, poorly timed, often ineffective
- Culture not tended; dysfunctional units

On the Folly of Rewarding A While Hoping for B

> "... reward systems that are fouled up in that the types of behavior rewarded are those which the rewarder is trying to discourage, while the behavior desired is not being rewarded at all."

Steven Kerr
Academy of Management Executive, 1995

On the Folly of Rewarding A While Hoping for B

> "… reward systems that are fouled up in that the types of behavior rewarded are those which the rewarder is trying to discourage, while the behavior desired is not being rewarded at all."

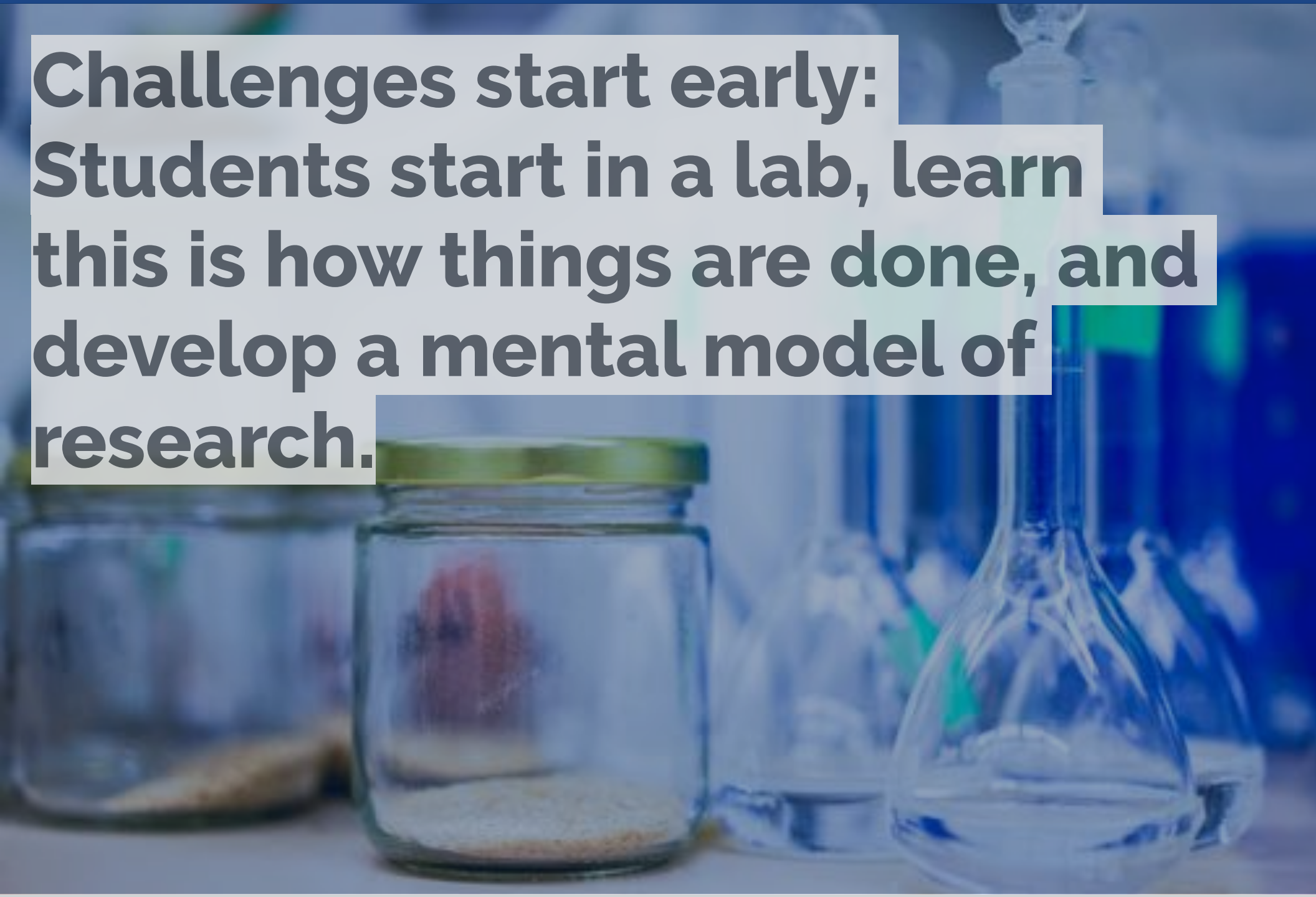
Steven Kerr
Academy of Management Executive, 1995

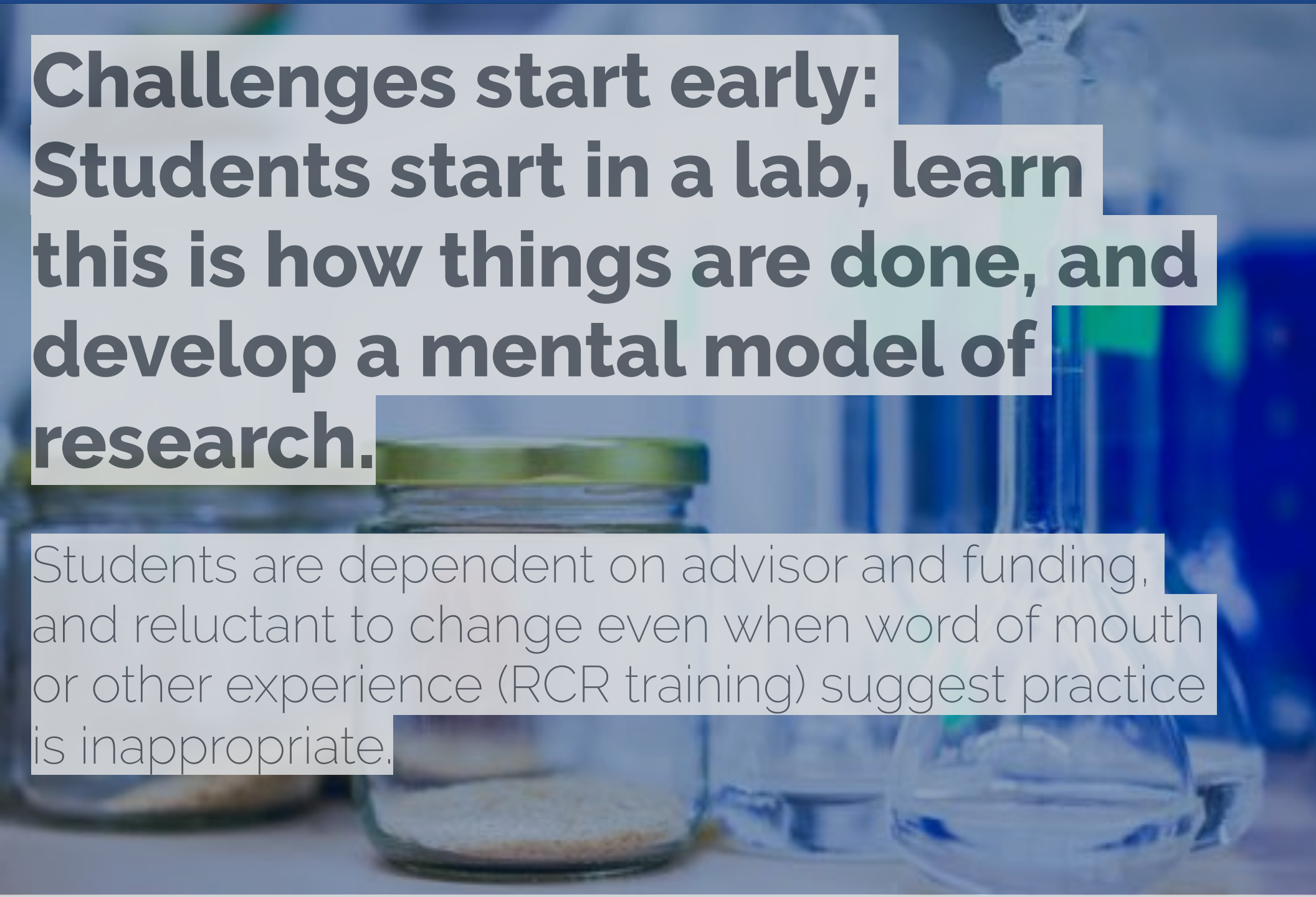
“The persistence of poor methods results partly from incentives that favour them, leading to the natural selection of bad science.”

—Smaldino and McElreath, 2016

Smaldino PE, McElreath R. 2016 The natural selection of bad science. R. Soc. open sci. 3: 160384. http://dx.doi.org/10.1098/rsos.160384

Bringing all these factors together for a wider view...


Challenges start early:
Students start in a lab, learn this is how things are done, and develop a mental model of research.



Challenges start early: Students start in a lab, learn this is how things are done, and develop a mental model of research.

Students are dependent on advisor and funding, and reluctant to change even when word of mouth or other experience (RCR training) suggest practice is inappropriate.

Career TRAGEDIES

- Temptation
- Rationalization
- Ambition
- Group, authority pressure
- Entitlement
- Deception
- Incrementalism
- Embarrassment
- Stupid Systems

Career TRAGEDIES

Temptation

Rationalization

Ambition

Group, authority pressure

Entitlement

Deception

Incrementalism

Embarrassment

Stupid Systems

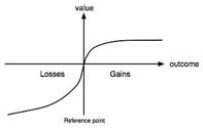
Example:

We can always justify improper actions to ourselves

Think about a graduate student who is about to submit a paper for publication that will determine job prospects.

- Experiments to complete
- Limited time to repeat and iterate
- Believes the research is good and important work
- Data *almost* tell the best story

Note: Students are even less likely to change if observed practices are “winning”



Loss Aversion +
Prevention Focus

Loss Aversion in Riskless Choice: A Reference-Dependent Model

Amos Tversky and
Daniel Kahneman

Loss Aversion

Losses loom larger than gains

Prevention Focus

People take more risks to avoid *losses*

Think about the mindset of a mid-career faculty member in the context of loss aversion and prevention focus

- People to support
- $ Grants to secure to keep it all going
- Promotion to full professor
- Papers required to do it all

These factors only intensify as scientists achieve greater professional success.

The Quarterly Journal of Economics
Vol. 106, No. 4 (Nov., 1991), pp. 1039-1061


CONTEXT
INDIVIDUAL

Focus on integrity mindset in research environments to reinforce rigor and reliability.

1

2

3

1

Measure it

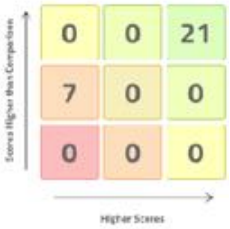
1

2002 IOM Report on Research Integrity

To promote responsible research conduct and fostering integrity**, institutions** should:

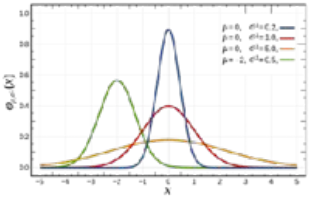
- Establish and **continuously measure** their structures, processes, policies, and procedures
- **Evaluate the institutional environment supporting integrity** in the conduct of research
- **Use this knowledge for ongoing improvement**

1

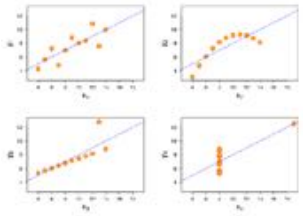


Survey of Organizational Research Climate (SOURCE)

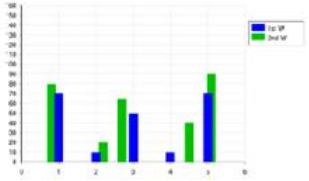
developed by Brian Martinson, Carol Thrush



Statistically validated with large sample



Scores show correlation between choices and research environment



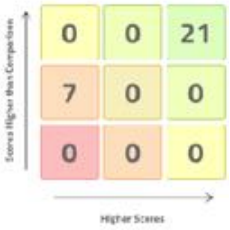
Benchmarking through two-stage reporting: campus and (anonymized) comparison database

Crain, A. Lauren, Brian C. Martinson, and Carol R. Thrush. 2013. "Relationships Between the Survey of Organizational Research Climate (SORC) and Self-Reported Research Practices." Science and Engineering Ethics 19 (3): 835–50.

1

Benchmarking power comes from competitive instincts of human beings. Let's harness that to improve.

1

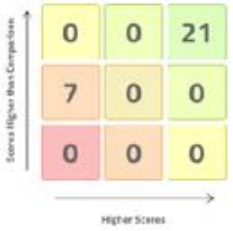


Survey of Organizational Research Climate (SOURCE)

- Automated data collection through emailed surveys
- Ability to measure success of efforts over time
- Full-service implementations, with consultation and repeat administration

NCPRE on-line engine and benchmark database

1



SOURCE

Putting it to work

Measure and assess the integrity of institutional research climates.

Find ways to influence them positively by studying the bright spots.

Give leadership tools for shaping environments

WORKPLACE CLIMATE

Metrics for ethics

Focus on perceived working conditions could help graduate schools to train responsible researchers.

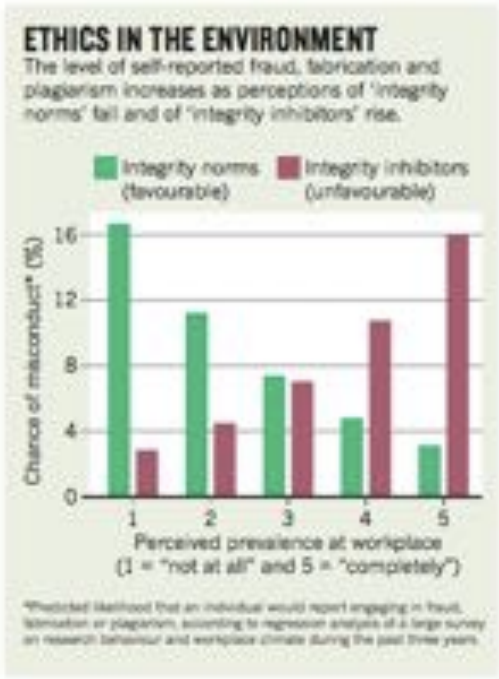
BY MONYA BAKER

Training in research ethics is mandatory for many US graduate students and postdocs, but there is little evidence that formal classes prompt scientists to conduct research ethically. However, the workplace climate — which includes perceptions of regulatory committees, data confidentiality and treatment of trainees — influences research practices and can spawn behaviours such as poor record-keeping or plagiarism.

An interdisciplinary team has developed a survey to assess work conditions in research institutions, with a long-term goal of establishing a baseline for measurements of workplace climate across disciplines and universities. The SOuRCe (Survey of Organizational Research Climate) is a 32-question survey that divides workplace climate into seven categories, including integrity norms (such as giving due credit to others' ideas), integrity inhibitors (such as inadequate access to material resources) and adviser–advisee relations. The team hopes that such data will help institutions to craft policies that will improve research conduct.

of respondents reported feeling ill-equipped to judge whether university policies support responsible research — which suggests that those topics are not discussed in meaningful ways, she says. Klomparens used the
spur faculty members in specific depa
talk to trainees about norms in autho
management and peer review. "Be
use the survey data by graduate pr
and by discipline, we can make reco
tions," she says. To encourage particip
emphasized to respondents that the
intended to shame or punish, and res
stripped of identifying information.

Brian Martinson studies research
at the non-profit HealthPartners In
Education and Research in Bloo
Minnesota, and helped to develop t
at 40 academic health centres (B. C. M
et al. Sci. Eng. Ethics **19**, 813–834;
has also worked on it in a separat
with MSU, Pennsylvania State Univ
the University of Wisconsin–Ma
poor workplace climate correlates w
undesirable research behaviours, eve
forms such as data falsification, he



2

Effective, realistic RCR

2

Research Ethics Programs

By some estimates, institutions devote less than 0.1% of research funding to RCR

Mostly delivered through on-line, multiple-choice programs (89.6% in one survey).

Because they are scalable, and documentable

Even that isn't reliably done

Focus is on rules and compliance vs. real problems encountered in research

2

One-size-fits-all multiple choice compliance training is not RCR.

2

Real-World Research Needs:

- Professional skills: present research, mentor, support diversity, good laboratory practices…
- How to have a dispute professionally
- How to maneuver in the trenches for getting credit and giving it vs. the formal rules of authorship
- How to choose a mentor and colleagues for character
- The line between making your data look "pretty" and manipulating/altering data and images
- Finding the line between inappropriate self-promotion and advancing your career sensibly
- How to get useful advice, and recognize it, when you encounter a problem

2

It Should Be:

- Relevant to the audience
- Required for all
- Interactive, experiential; using best practices
- Meaningful: related to work being done
- Delivered at least in part by respected researchers
- Assessed

3

And, what about that barrel?

3

Some Factors

Institutional leadership, structures

Reward Systems, institutional and individual

Conflicts of Interest


3
Tone
Mood
Buzz

3

Why don't we hold leadership accountable when culture does not support rigor, investigations are botched?

3

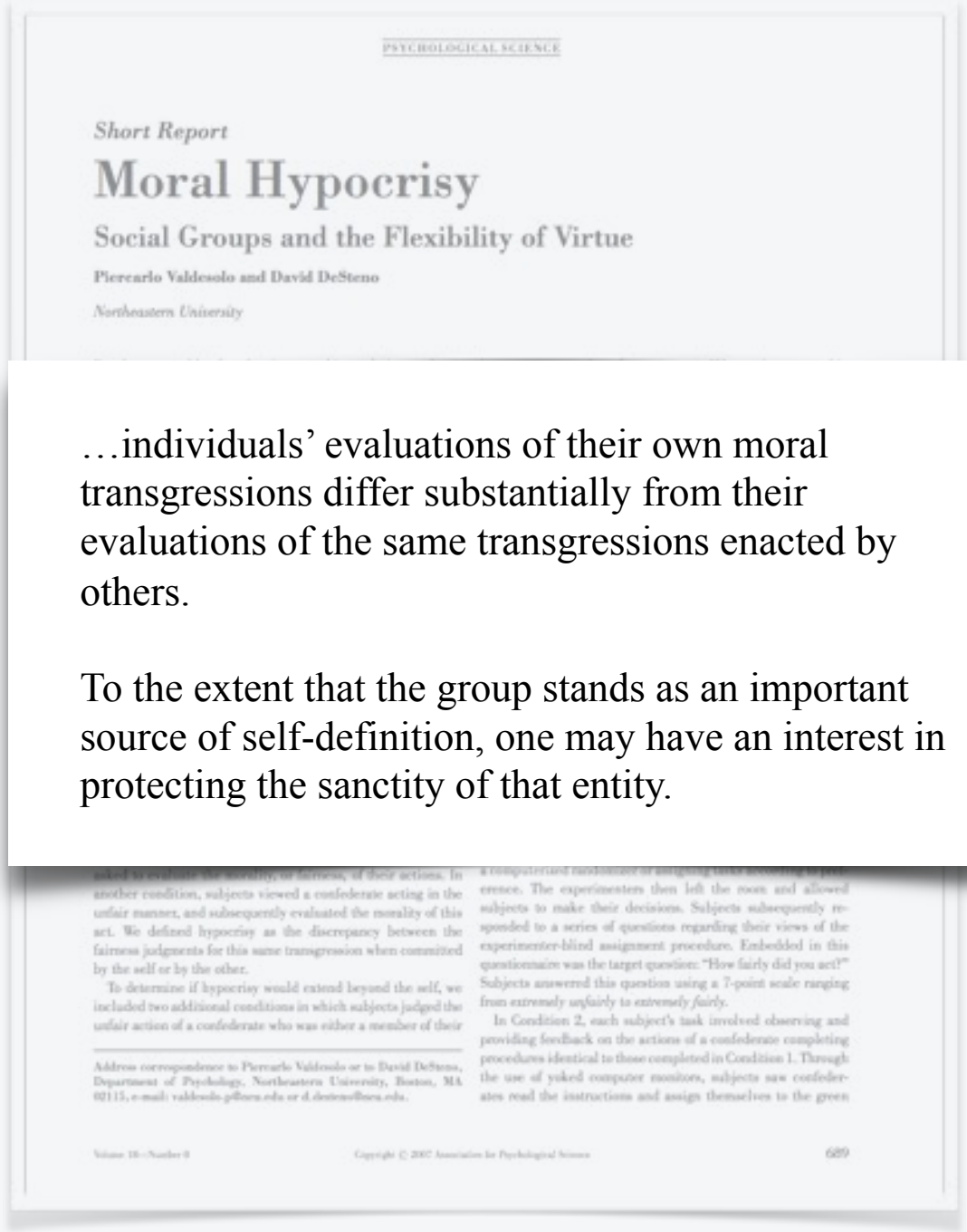
Institutional response to problems:

Circling the wagons

Wr⊘ng Questions!

1. "How will this affect our reputation if it becomes known?"
2. "How could anyone think Bill would do such a thing?"
3. "How can we make this go away?"
4. "We don't have to report this, do we?"
5. "Why would you want to cause trouble for your own research project?"

3



PSYCHOLOGICAL SCIENCE

Short Report

Moral Hypocrisy

Social Groups and the Flexibility of Virtue

Piercarlo Valdesolo and David DeSteno

Northeastern University

…individuals' evaluations of their own moral transgressions differ substantially from their evaluations of the same transgressions enacted by others.

To the extent that the group stands as an important source of self-definition, one may have an interest in protecting the sanctity of that entity.

3

PSYCHOLOGICAL SCIENCE

Short Report

Moral Hypocrisy

Social Groups and the Flexibility of Virtue

Piercarlo Valdesolo and David DeSteno

Northeastern University

…individuals' evaluations of their own moral transgressions differ substantially from their evaluations of the same transgressions enacted by others.

To the extent that the group stands as an important source of self-definition, one may have an interest in protecting the sanctity of that entity.

The New York Times | https://nyti.ms/2kBv1wt

PRO FOOTBALL

Why Do Fans Excuse the Patriots' Cheating Past?

Sports of The Times

By JULIET MACUR FEB. 5, 2017

BOSTON — In a psychological experiment, researchers separated people into two groups and offered some of them an option: Complete a fun, 10-minute task, or take on a difficult, 45-minute one. Placed in a room alone, they were told to choose

3

Bad Practices in University Reports Seen by NSF Office of Inspector General

snippet is not plagiarism ergo totality is not plagiarism

- Asking leading questions to allow subject to explain a way out instead of asking more pointed questions like “Did you do it?”
- Investigative report lacks supporting evidence and fails to adequately address the elements of a research misconduct finding
- Convene first committee meeting on day 175 (due to OIG by day 180)
- Half page investigation report
- Pre-written admission for grad student; sign or we investigate
- Fail to interview key witnesses

3

Investigation Shortcomings

Inadequate reports

Missing elements

Poor record keeping

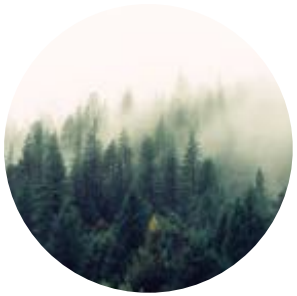
Evidence ignored

Missed and inadequate interviews

Wrong standards of proof

Misunderstanding "intent"

Wrong definitions



Wrong perspective

Ghost investigation

3

Potti found guilty of research misconduct

By Jocelyn Kaiser | Nov. 9, 2015 , 12:30 PM

"We are pleased with the finding of research misconduct by the federal Office of Research Integrity related to work done by Dr Anil Potti. We trust this will serve to **fully absolve the clinicians and researchers who were unwittingly associated with his actions, and bring closure to others who were affected"**

[institutional representative]

deposition

Q: "Once you started digging, how long did it take you to find the manipulations that had been done?"

A: "It would take you maybe an hour."

To date: 11 retractions, 64 co-authors

3 clinical trials, 117 patients

Better Questions

1. "Do we want our names and reputations associated with an institution where dishonest work is countenanced?"

2. "Are there other scholars depending on this work?"

3. "What kind of education are students getting at our institution?"

Peer Review Institutional Investigation Plans, Reports

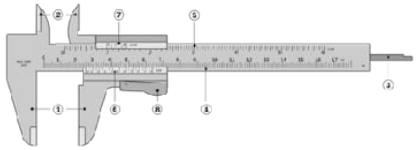
- Does the investigation plan identify the right questions and propose a meaningful approach?
- Were the correct people interviewed? All of them?
- Were the relevant data reviewed by appropriate experts?
- Does the investigation report provide factual basis and data?
- Are the conclusions of the report clearly supported?

1992

"The argument that science must be regulating itself pretty well because it is making progress is far from compelling; perhaps progress would be twice or four times as fast with greater 'scrupulousity.'"

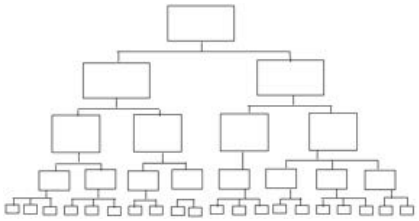
On Misunderstanding Scientific Misconduct
Paul J. Friedman
Knowledge: Creation, Diffusion, Utilization.
vol. 14 No. 2, December 1992 153-156

Let's ask more questions:



Why aren't we using an empirical approach?
Why aren't we assessing our environments?
Why aren't we using results to improve ?

Why is doing RCR well such a low-priority?
Why are only students required to take RCR?
Why aren't professional and real-world skills included in RCR?



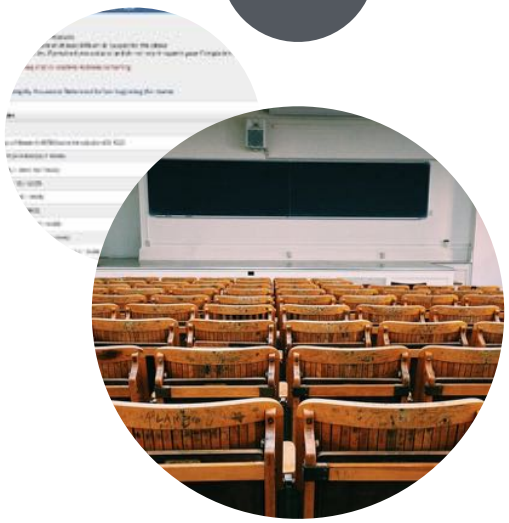
Why aren't we reforming perverse incentives?
Why aren't institutional leaders queried about the integrity of their environments?
Why aren't investigation reports peer reviewed?
Why aren't leaders who preside over botched investigations held accountable?

Let's get serious about modifying our environments, systems, and practices to reinforce an integrity mindset.

1

Assess and benchmark

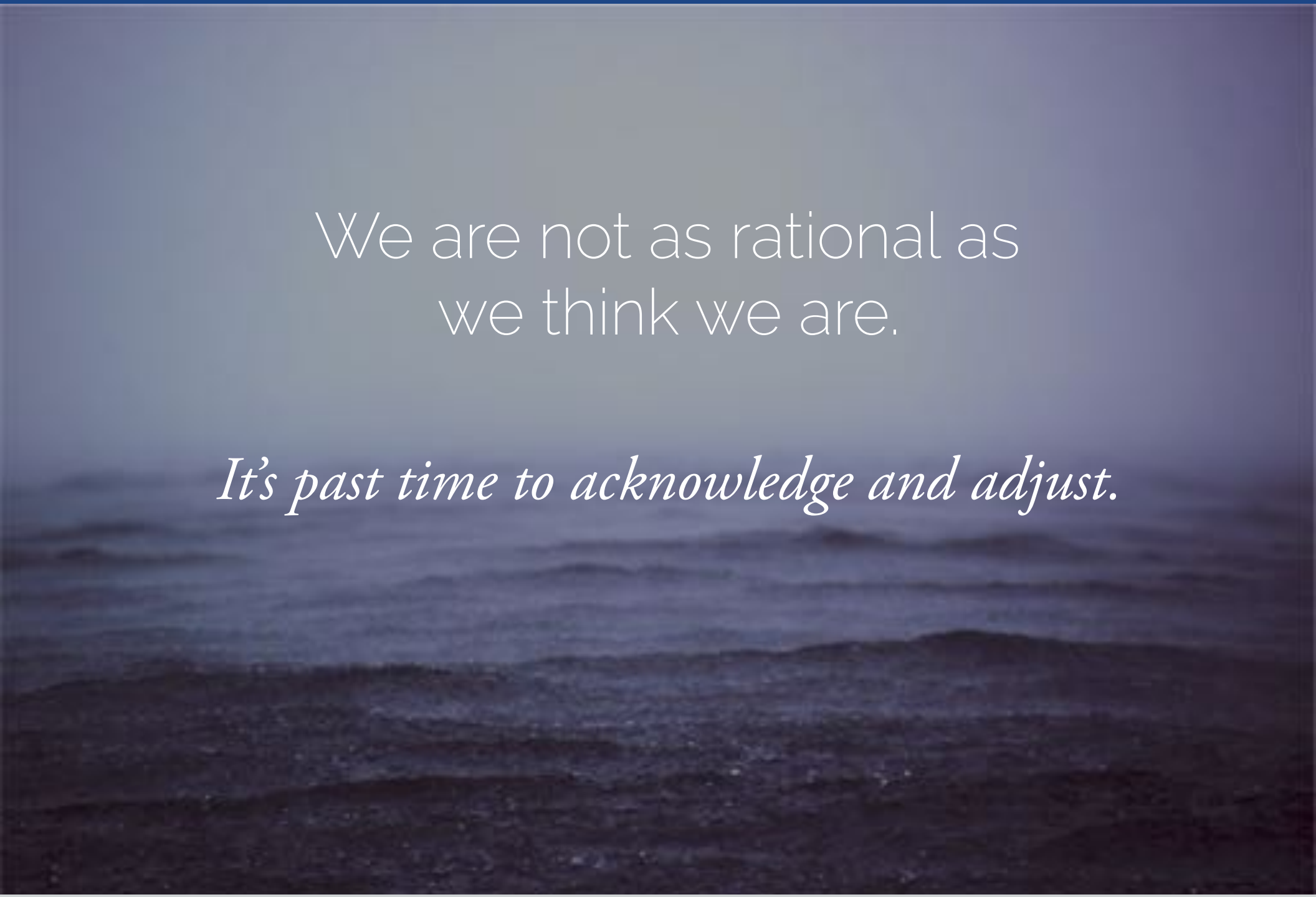
2



Better RCR

3

Improve institutional stewardship


We are not as rational as
we think we are.
It's past time to acknowledge and adjust.