Taylor & Francis
Taylor & Francis Group

Understanding Research Misconduct: A Comparative Analysis of 120 Cases of Professional Wrongdoing

James M. DuBois, D.Sc., Ph.D.,[1] Emily E. Anderson, Ph.D., M.P.H.,[2] John Chibnall, Ph.D.,[3] Kelly Carroll, J.D., M.B.E.,[4] Tyler Gibb, J.D.,[4] Chiji Ogbuka, M.A., M.Div.,[4] and Timothy Rubbelke, B.S.[4]

[1]Bander Center for Medical Business Ethics, Saint Louis University, St. Louis, Missouri, USA
[2]Neiswanger Institute for Bioethics, Loyola University of Chicago, Chicago, Illinois, USA
[3]Department of Neurology and Psychiatry, Saint Louis University, St. Louis, Missouri, USA
[4]Albert Gnaegi Center for Health Care Ethics, Saint Louis University, St. Louis, Missouri, USA

We analyzed 40 cases of falsification, fabrication, or plagiarism (FFP), comparing them to other types of wrongdoing in research (n = 40) and medicine (n = 40). Fifty-one variables were coded from an average of 29 news or investigative reports per case. Financial incentives, oversight failures, and seniority correlate significantly with more serious instances of FFP. However, most environmental variables were nearly absent from cases of FFP and none were more strongly present in cases of FFP than in other types of wrongdoing. Qualitative data suggest FFP involves thinking errors, poor coping with research pressures, and inadequate oversight. We offer recommendations for education, institutional investigations, policy, and further research.

Keywords: fabrication, falsification, historiometry, plagiarism, professional misbehavior, research misconduct

Address correspondence to James M. DuBois, D.Sc., Ph.D., Bander Center for Medical Business Ethics, Salus Building 5th Floor, 3545 Lafayette Ave, St. Louis, MO 63104, USA. E-mail: duboisjm@slu.edu

INTRODUCTION

Research and experience tell us that people find things interesting when they are new, unexpected, and complex (Silvia, 2005). Discussing why research misconduct is wrong meets none of these criteria. Falsification and fabrication of data constitute a form of lying and plagiarism a kind of stealing. As children most of us learned that lying and stealing are wrong. It is also relatively uninteresting to discuss the negative consequences of research misconduct because their harmful effects are so immediately evident: Fabrication and falsification contaminate scientific literature with incorrect information, which wastes funds, hampers medical and technological development, and poses risks to patients and consumers, while plagiarism deprives authors of credit for their work.

However, we propose that the question *"why do investigators engage in research misconduct?"* is interesting because misconduct cases are typically shrouded in secrecy, and studies on professional wrongdoing suggest that the causes of misconduct are complex. These characteristics of research misconduct—secrecy and complexity—also make it very difficult to study the phenomenon.

THE PROBLEM OF SECRECY

A meta-analysis of surveys on research misconduct found that almost 2% of investigators surveyed admitted to fabricating or falsifying data at least once; when asked about the behavior of their colleagues, approximately 14% reported observing a case (Fanelli, 2009). Our team's recent survey of the 194 comprehensive doctoral or medical schools in the United States found that 88% of research integrity officers had investigated a credible case of wrongdoing over the past 2 years (range 1–16 or more cases); 54% of these cases involved research misconduct (DuBois et al., 2013a).

With such prevalence rates, one would expect that research misconduct would be well studied and well understood. However, most cases are not reported to federal authorities—even cases involving federal funding (Titus et al., 2008). Cases that are reported are investigated confidentially. Although findings are published, federal investigators currently retain little detailed information on wrongdoers' reported motives or their research environments. When investigation reports are provided in response to a Freedom of Information Act request, the quality and quantity of information is highly varied depending on the quality of the institutional investigation and degree of redacting deemed appropriate (personal communication with John Dahlberg, Office of Research Integrity, January 29, 2013). Even when an investigation finds the accused to be guilty of misconduct, our extensive literature reviews over the past four years have found that most such cases receive scant attention from investigative reporters.

THE PROBLEM OF COMPLEXITY

Wrongdoing in research is complex in at least two important ways. First, wrongdoing in research takes many different forms (DuBois et al., 2012; Martinson et al., 2005). In "Phase I" of our project, described in detail elsewhere (DuBois et al., In press), we researched and wrote synopses of 100 highly heterogeneous cases of wrongdoing in healthcare delivery and research. Five experts rated the severity or seriousness of wrongdoing found in each case, and we identified variables that correlated with more severe forms of wrongdoing. One lesson learned was that wrongdoing is not a homogeneous phenomenon; different kinds of wrongdoing are associated with different variables. Thus, it is misleading to speak of "organizational deviance" or "professional misbehavior" as though it is a unified phenomenon. Accordingly, we need to inquire separately into different forms of wrongdoing:

1. Why would someone fabricate research data?

2. Why would a clinical investigator enroll patients who do not meet the inclusion criteria approved by an institutional review board (IRB)?

3. Why would an animal researcher not follow a protocol that requires administration of anesthesia before performing painful procedures?

Second, there are many reasons why a researcher might behave unprofessionally. Some reasons concern personality traits. For example, an overly robust sense of self-entitlement and cynicism have been associated with poor ethical decision making (Mumford et al. 2006). Other explanations concern individual psychological responses to environmental factors. For example, several studies have found that stress—whether cognitive overload or exposure to stressful life events—is associated with diminished ethical decision-making (Mumford et al., 2001) and research misconduct (Davis et al., 2007). Other researchers have found that feeling unfairly treated by a system may contribute to skirting rules within that system (Martinson et al., 2006; Martinson et al., 2010; Keith-Spiegel and Koocher, 2005). Davis (2003) hypothesizes that culture may play a role in research misconduct: cultural norms inculcated in one's nation of birth or training may contribute to behavior that is deviant within U.S. culture.

Finally, a review of recent criminological studies suggests that environmental factors that create opportunity for research misconduct deserve closer attention (Adams and Pimple, 2005). Consistent with this view, our team conducted a systematic literature review and identified ten environmental factors that may predict wrongdoing in research:

1. Playing conflicting professional roles (Grover, 1993; Levine, 1992);

2. Having financial rewards for wrongdoing (Jennings, 2006; Rodwin, 1993);

3. Others benefitting from wrongdoing (Victor et al., 1993);

4. Being penalized for following the rules (Hegarty and Sims, 1978; Jansen and Von Glinow, 1985);

5. Others being penalized for doing what is right (e.g., reporting wrongdoing) (Schuchman, 2008);

6. Unjust treatment of employees or unfair processes (Greenberg, 1993; Keith-Spiegel and Koocher, 2005; Martinson et al., 2006);

7. Ambiguous legal or ethical norms (Davis, 2003; Meyer and Bernier, 2002; Shah et al., 2004);

8. Vulnerable victims (Bandura, 1999; Zimbardo, 2007);

9. Oversight failures;

10. Occupying a position of authority over peers (Bramstedt and Kassimatis, 2004; Marshall, 1999).

As we show in this previously published review, each of these ten factors can be understood as contributing motive, means, or opportunity (or a combination thereof) for wrongdoing in research, each is supported by some empirical studies, and each is evidenced in one or more cases our team has studied (DuBois et al., 2012).

RATIONALE FOR A HISTORIOMETRIC APPROACH

Historiometry can be defined as the statistical analysis of coded data from historical narratives. Our use of the method involves identifying appropriate cases, reading all the published literature we can find on a case, coding the presence of more than 50 variables, producing descriptive statistics, and analyzing data to identify which variables are significantly associated with a specific kind of wrongdoing. The limitations of such an approach are fairly obvious. For example, it is difficult to control for confounding variables when examining non-randomly assigned variables, and published reports may omit important information. However, historiometry offers several major advantages over more common methodologies: It can examine the relationship of a large number of variables within one study; its data sources foster ecological validity; and it is feasible.

Multivariable Friendly

Given the rich array of theories and data on the causes of research misconduct, it is important to identify a method that can track the influence of more than one variable at a time. Yet most experimental designs test the influence

of a very small set of independent variables. In contrast, historiometry allows us to extract numerous variables from published reports and build complex statistical models (limited primarily by sample sizes rather than participant time).

Ecological Validity

Social psychology abounds with rigorous, prospective, experimental data on ethical decision-making obtained from university students with no history of wrongdoing. It is unclear whether such data—despite its methodological elegance—generalizes well to real world cases of wrongdoing among established researchers. Apart from differences in the relevant populations, the recent movement toward effectiveness research rests on the observation that randomized controlled trials often leave much to be desired when it comes to knowledge of real world applications (Depp and Lebowitz, 2007).

We have very little data obtained directly from those who have engaged in wrongdoing, and much of what we do have derives from self-reports from individuals who may have very little insight into social factors that contributed to their behavior: Data suggest that self-serving biases operate below the level of awareness, and individuals may easily ignore the ways that system failures contributed to their behavior (Dana and Loewenstein, 2003; Katz et al., 2003). Historiometric data—being drawn from real world cases—has a much greater chance of being ecologically valid than alternative designs (Brewer, 2000).

Feasibility

It is not likely that anyone could conduct a study of research misconduct that involves a large random sample. Falsification, fabrication and plagiarism are relatively rare events that occur in the privacy of offices and laboratories. They are also events that we do not want to induce in a prospective manner, for example, through random manipulation of variables that may predict misconduct because the consequences are so serious. In contrast, historiometric studies are feasible.

Rationale for a Comparative Approach

As noted above, we cannot treat wrongdoing in research as though it is a unified phenomenon such as "organizational deviance." Accordingly, in this article we examine three different sets of relatively homogeneous cases: research misconduct, other wrongdoing in research, and wrongdoing in medical practice. Although many cases of professional misbehavior involve multiple forms of wrongdoing, cluster analysis of our first 100 cases indicated that research misconduct does not overlap with human subjects or animal care

violations or with violations in the domain of medical practice. This solves a major problem for the historiometric study of research misconduct, namely, that no one publishes detailed accounts of "good research," which could serve as comparison or baseline cases.

METHODS

Our method involved identifying appropriate cases; conducting literature reviews on cases; reading all relevant literature; extracting and rating the presence of key variables; and statistically analyzing data to produce descriptive statistics and to examine the relationship between research misconduct and independent variables. Each element of our method is described below.

Case Sampling

As is common in historiometric research (Simonton, 2003; Mumford, 2006), we used criterion-based sampling. Table 1 presents our sampling criteria with a rationale for each. To identify which variables characterize research misconduct in a statistically significant manner, we sampled three distinct kinds of wrongdoing: research misconduct (FFP), other research wrongdoing (such as

Table 1: Case Inclusion Criteria

Criterion	Rationale
Location: USA	- Provides a relatively clear and stable set of laws and cultural norms surrounding the 3 sampling areas
Date of occurrence: Jan. 1980–Dec. 2011	- Case conclusion must be at least 1 year old to allow adequate reporting of information on the case - Because widely published cases of misconduct are rare, it is important to work with a broad range of dates - By 1980 norms prohibited all of the behaviors under investigation
Kind of wrongdoing: (1) misconduct, (2) other research violations, or (3) wrongdoing in medical practice	- All involve serious professional violations - Cases are distinct and may serve as comparison cases
Case coverage: Case is described in at least 5 published sources, including articles and reports of agencies and boards	- Indicates social significance of case - Ensures adequate availability of information on the case
Case involves a clear individual wrongdoer upon whom we can focus	- We aim to understand and predict individual behavior, rather than corporate behavior, which may have different dynamics

violations of animal care or of human subjects protections), and wrongdoing in medical practice (such as boundary violations, abuse of prescribing privileges, or fraudulent unnecessary procedures). All kinds of wrongdoing represented in our sample appear in a taxonomy we developed to guide sampling; the taxonomy has been used by our team with a free-marginal multirater Kappa coefficient of .85 (DuBois et al., 2012).

We decided to sample 40 cases of each kind of wrongdoing to create samples of equal size that, based on effect sizes observed in our past studies (DuBois et al., 2013b), would be large enough to detect significant differences if such differences existed.

Once we established our sampling criteria, we conducted extensive reviews of the literature. Many cases were excluded because they were not described in a minimum of 5 publications. On the one hand, this means that our cases may not be representative of all such cases occurring in the United States during our sampling timeframe; on the other hand, Simonton argues that historiometric methodology does well to focus on the most well-known, or in our study, notorious cases: "The most eminent individuals in a domain are not only the most representative of the phenomenon of interest, but information about such subjects is likely to be more extensive and reliable" (Simonton, 2003, p. 625).

Our literature reviews were conducted by developing search criteria in consultation with a research librarian and using search terms in LexisNexis Law (searching federal and state cases, newspapers, and magazines), Medline, Google, and Google Scholar. Additionally, for misconduct cases we examined all findings published by the Office of Research Integrity (ORI); for cases of other research violations we examined all findings published by the Office of Human Research Protections; and for cases of wrongdoing in medical practice we examined reports from the HHS Office of Inspector General and state medical boards.

Data Extraction

A research assistant (RA) read all publications relevant to a case (an average of 29 articles per case). Using a case datasheet and scoring guide, RAs extracted data on 51 variables pertaining to the setting (n = 5), the wrongdoer (n = 8), the wrongdoer's history of misbehavior and illegal activity (n = 4), how the case was reported and investigated (n = 7), consequences to the wrongdoer (n = 6) and wrongdoer's institution (n = 3), environmental variables (n = 12, scale items), case duration, complexity, and media coverage (n = 3), and the RA's theory of the case, including notable social dynamics and wrongdoer motives (n = 3, open-ended). While our primary independent variables of interest are the 10 environmental factors enumerated in the introduction, many of these variables attempt to address at least indirectly some of the individual

variables mentioned in the introduction. For example, we tracked the individual's nation of origin and training to take into account the possible role of culture, and we tracked personal problems such as divorce or declarations of bankruptcy as markers of stressors.

Continuous variables (such age, date, number of sources consulted) were entered as such; bivariate variables (such as foreign-trained or evidence of insanity plea) were entered as yes/no; and scaled items were scored on a scale from 1 to 3 using a benchmark scoring guide described in DuBois et al. (2013b). The meaning of the scale was defined for each variable, e.g., "1 = no evidence of conflicting roles; 2 = conflicting roles were played but managed; and 3 = conflicting roles were played and not managed." The guide provides examples of each score for each variable.

The team ensured the quality of case research through several means: RAs were all Ph.D. students in health care ethics who read a series of 21 articles explaining the relevance of each primary variable we tracked and then received 8 hours of formal training prior to researching their first case; individual RA's researched anywhere from 7 to 24 cases across the life of the project. All cases were reviewed by a Ph.D. co-investigator who read at least three articles on each case (Anderson); and the team met weekly with the principal investigator (DuBois) to discuss case research, questions on the scoring guide, and procedures.

For this dataset only one RA per case produced all ratings because we previously established very high inter-rater reliabilities (ICC = .84–1.0) for all scaled values using a set of 100 cases (DuBois et al., 2013b).

Data Analysis

Data analysis involved generating descriptive statistics (frequencies and means) and analyzing the relationship of the kind of wrongdoing to all independent variables using Chi square to test for differences in frequency and ANOVA to test for differences in mean values on scaled items.

RESULTS

In what follows, we present descriptive statistics for the three forms of wrongdoing with tests of significance aimed at identifying when a given variable is more strongly associated with a specific kind of wrongdoing. We use tables to present statistical details; our text summarizes results. We conclude with a presentation of qualitative data on wrongdoer motives and social dynamics.

Frequency of Wrongdoer and Setting Variables

Table 2 presents frequencies for wrongdoer and setting variables. Chi square (X^2) values are presented to indicate when the frequency of a variable (such as being male or being trained abroad) differs significantly across kinds of wrongdoing.

Highly publicized cases of research misconduct are characterized by very few setting traits: They take place in academic medical centers (90%) with government funding (70%). However, this is likely because non-government funded cases need not be reported to oversight agencies and can be handled relatively quietly. Whereas ORI and NSF publish outcomes of their investigations, private industry and universities do not.

Most individuals who engage in research misconduct are male (75%), but the rate is not higher than the prevalence of males receiving NIH research grants (82%) at the median time of the cases (Office of Statistical Analysis and Reporting, 2012).

The individuals' history of wrongdoing is not as pronounced in research misconduct cases as in other kinds of cases: Most individuals repeated their misconduct prior to being investigated formally (68%); but fewer than half engaged in different kinds of wrongdoing (43%) or committed violations at multiple institutions (18%).

Table 2: Wrongdoer and Setting Variables: Comparison of Frequencies

Independent Variable	Research Misconduct	Other Research	Medical Practice	X^2
Setting variables				
Academic medical setting	90%$_a$	75%$_a$	3%$_b$	71.03***
Government funding	70%$_a$	58%$_a$	0%$_b$	45.63***
Private funding	33%$_{a,b}$	63%$_a$	0%$_b$	36.12***
Had an accomplice	10%$_a$	40%$_b$	35%$_b$	10.18**
Others found guilty	20%	33%	35%	2.50
Wrongdoer Variables				
Male	75%$_a$	90%$_b$	93%$_b$	5.89*
Born outside United States	10%	20%	20%	1.92
Trained outside United States	15%	18%	25%	1.40
Plea of insanity	0%	0%	5%	4.07
Found unfit to stand trial	0%	0%	0%	–
Evidence of addiction	0%$_a$	0%$_a$	10%$_b$	8.28*
Significant personal problems	5%$_a$	0%$_a$	23%$_b$	13.41***
Claimed following orders/policy	3%	10%	0%	5.43
History of Wrongdoing				
Wrongdoing was repeated	68%$_a$	80%$_{a,b}$	95%$_b$	9.79**
Different kinds of wrongdoing	43%$_a$	65%$_{a,b}$	75%$_b$	9.30**
Wrongdoing in multiple institutions	18%	20%	28%	1.28
Felony arrests in personal life	3%$_a$	0%$_a$	13%$_b$	7.37*

* = $p < .05$, ** = $p < .01$, *** = $p < .001$. N = 120 cases: 40 Research Misconduct, 40 Other Research, 40 Medical Practice. When chi-squared is significant, percentages that do not share subscripts are significantly different by standardized residual.

Five percent or fewer attributed their research misconduct to significant personal problems, addictions, or mental disorders, and none had been convicted of felonies in their personal lives; most of these factors are significantly more likely to be found in cases of medical practice violations.

Individuals who are born or trained outside of the United States are not represented in our sample of highly publicized cases of research misconduct with any greater frequency than in other kinds of wrongdoing, and they are represented in proportion to the percentage working in the field around the median date in our study (Davis, 2003).

Research misconduct cases involve highly focused kinds of wrongdoing; to the extent they involve another wrongdoing, it is usually an extension of the primary wrongdoing—e.g., a violation of publication ethics.

Frequency of Reporting, Investigation, and Consequences Variables

Table 3 presents frequencies and Chi square results on the reporting (whistleblowing), investigation, and consequences of cases. In 28% of cases there was a failed attempt at reporting research misconduct (that is, the wrongdoing continued for some time following an initial report); however, this did not occur at a greater rate than in other kinds of wrongdoing.

Cases of research misconduct are far more likely to have an institutional whistleblower than other research wrongdoing or wrongdoing in medical practice, with subordinates comprising the largest group of whistleblowers (23%).

Not surprisingly, most cases involved an institutional (88%) and federal investigation (85%)—rates significantly higher than found in cases involving other kinds of wrongdoing. Consequences to the wrongdoer were also more severe, typically including loss of job or funding opportunities (90%).

Environmental Variables

Very few of the variables that were hypothesized to predict professional misbehavior were present in cases of research misconduct. Only two variables were rated higher than 1.5 on a scale from 1 to 3: Oversight failures, which in the context of research misconduct can include failures of research integrity officers to investigate a case in a timely manner or of collaborators to investigate suspicious data (m = 1.65) and position of authority (m = 1.90). However, oversight failures play less of a role in cases of research misconduct than other kinds of wrongdoing and the prevalence of individuals in positions of authority was likely due to the fact that principal investigators are ordinarily held accountable for the integrity of data published in papers funded by their grants.

Table 3: Reporting, Investigation, and Consequences Variables: Comparison of Frequencies

Independent Variables	Research Misconduct	Other Research	Medical Practice	X^2
Failed reporting attempts	28%	35%	43%	1.98
Whistleblower Description+				
Patient/participant/family	$0\%_a$	$18\%_{a,b}$	$45\%_b$	38.36***
Subordinate	$23\%_a$	$10\%_{a,b}$	$5\%_b$	
Peer: institutional	$15\%_a$	$3\%_b$	$3\%_b$	
Peer: external	5%	8%	8%	
Oversight personnel: institutional	$15\%_a$	$8\%_{a,b}$	$3\%_b$	
Oversight personnel: external	5%	13%	10%	
Others (e.g., reporter)	18%	25%	15%	
Unknown	20%	18%	13%	
Investigation				
Institutional investigation	$88\%_a$	$53\%_{a,b}$	$25\%_b$	31.72***
Board or professional body investigation	$20\%_a$	$23\%_a$	$85\%_b$	44.40***
Regulatory oversight body investigation	$85\%_a$	$75\%_a$	$53\%_b$	10.73**
Criminal investigation	$33\%_a$	$20\%_a$	$73\%_b$	24.75***
Civil investigation	$10\%_a$	$43\%_{a,b}$	$60\%_b$	21.97***
Consequences to Wrongdoer				
Loss of job, opportunities, funding . . .	$90\%_a$	$50\%_b$	$85\%_a$	20.27***
Loss of licensure/credentialing	$20\%_a$	$15\%_a$	$85\%_b$	50.83***
Financial penalties	$33\%_a$	$30\%_a$	$75\%_b$	20.61***
Prison, probation, house arrest	$25\%_{a,b}$	$8\%_a$	$65\%_b$	31.68***
Treatment, rehabilitation, education	18%	13%	30%	4.06
Wrongdoer still working in the field	$48\%_{a,b}$	$25\%_a$	$78\%_b$	22.20***
Consequences to Institution				
Publication of oversight failure	$33\%_a$	$70\%_b$	$28\%_a$	17.58***
Financial penalties, fines, settlements	18%	30%	18%	2.45
Impact on mission/increased audits . . .	$13\%_a$	$48\%_b$	$13\%_a$	17.82***

* = $p < .05$, ** = $p < .01$, *** = $p < .001$. N = 120 cases: 40 Research Misconduct, 40 Other Research, 40 Medical Practice. When chi-squared is significant, percentages that do not share subscripts are significantly different by standardized residual.
+ Whistleblower role was run as 1 variable with 8 roles, thus one X^2 value is reported for all roles.

Two variables were nearly absent across all kinds of misbehavior: wrong-doers who were reacting to organizational injustice, and wrongdoers who were penalized for doing what is right. All other variables were nearly absent in cases of research misconduct, but more strongly present in other areas of wrongdoing, including especially ambiguous norms, vulnerable victims, conflicting roles, and financial rewards.

We may speak of cases that are more "serious" insofar as they last longer and involve multiple kinds of wrongdoing. Several variables correlated with more severe cases of misconduct. Financial rewards to the wrongdoer (rho = .55, $p < .01$) and to others involved in the case (rho = .57, $p < .01$), as well as oversight failures (rho = .32, $p < .05$), were correlated with an increased number of violations. The duration of a case was correlated with the seniority (authority level) of the investigator (rho = .43, $p < .01$) with cases

Table 4: Environmental Variables: Comparison of Mean Scores (ANOVA)

Environmental Variables	Research Misconduct[+]	Other Research	Medical Practice	F
Conflicting roles++	1.15$_a$ (.53)	1.75$_b$ (.95)	1.05$_a$ (.32)	13.28***
Financial reward++	1.33$_a$ (.69)	2.00$_b$ (.93)	2.23$_b$ (.95)	11.70***
Others benefit++	1.30$_a$ (.61)	1.98$_b$ (.80)	1.83$_b$ (.90)	8.26***
Penalized for doing right	1.00 (.00)	1.00 (.00)	1.02 (.16)	n/a
Others penalized++	1.30 (.65)	1.18 (.55)	1.20 (.46)	.56
Mistreatment of wrongdoer	1.08 (.27)	1.02 (.16)	1.00 (.00)	n/a
Ambiguous norms++	1.02$_a$ (.15)	1.43$_b$ (.55)	1.05$_a$ (.22)	16.04***
Vulnerable victims++	1.07$_a$ (.35)	1.70$_b$ (.79)	1.68$_b$ (.69)	12.22***
Oversight failures++	1.65$_b$ (.77)	2.13$_b$ (.79)	1.78$_{ab}$ (.77)	4.03*
Position of authority++	1.90$_a$ (.74)	2.45$_b$ (.68)	1.55$_b$ (.81)	14.73***
Collaboration	1.20 (.46)	1.33 (.47)	1.00 (.00)	n/a
Other Variables				
History of wrongdoing++	1.30$_a$ (.88)	1.65$_{ab}$ (.86)	2.10$_b$ (1.03)	7.45***
Duration	3.83 (1.28)	4.30 (1.47)	4.22 (1.48)	1.31
Number of violations++	2.03$_a$ (1.48)	3.63$_b$ (1.64)	3.13$_b$ (1.49)	11.33***
Number of Sources	19.30$_a$ (17.52)	32.23$_b$ (23.98)	34.38$_b$ (31.19)	4.30*

* = p < .05, ** = p < .01, *** = p < .001. N = 120 cases; 40 Research Misconduct, 40 Other Research, 40 Medical Practice. When ANOVA is significant, means that do not share subscripts are significantly different by Tukey post hoc pairwise comparison.
+ Scores are means with standard deviations in parentheses.
++ MANOVA including these 10 variables was statistically significant, Wilks' Lambda = .309, p < .001.

involving more senior investigators lasting longer. These are all moderately strong correlations.

Across all categories of wrongdoing, cases last a surprisingly long time—between 3.8 and 4.2 years.

Qualitative Theory

In contrast to cases of other kinds of wrongdoing, very little seems distinctive about the research misconduct cases. A wide variety of environmental and individual factors that are frequently hypothesized to correlate with professional misbehavior were nearly absent from our dataset of highly publicized cases; the remaining variables were only moderately present. So how is misconduct to be explained? Drawing upon their extensive reading on the cases, RAs were asked to respond to three qualitative questions that would allow us to build a theory of each case. They typically wrote 3–5 sentences in response to each question. The following are the questions with the instructions provided to RAs:

1. **Describe Notable Social Dynamics**. Insofar as they are observed, note at least the following things: highly competitive environment; high profile work; high pressure, high quotas; prominent non-financial rewards; modeling of bad behavior by peers or supervisors; generally poor moral climate.

2. **Describe the Wrongdoer's Intentions and Character**. Describe the wrongdoer's intentions as you understand them, and his/her response to the wrongdoing. Address whether she/he expressed remorse or apologized. Did he/she lie or try to cover up wrongdoing? Blame others? You may also address personality factors mentioned by in the literature.

3. **What is Your Theory of the Case?** Provide your common sense theory of how the case occurred, using the theoretical framework embraced by this project (the interaction of individual and environmental factors in providing motive, means or opportunity). You may want to note when desirable information is missing, e.g., "no motive is apparent."

We used an open-ended format for these items in part because we believed that it would be difficult to quantify accurately variables such as the degree of competitiveness in an environment or the investigator's intentions given our data sources, which do not consistently address such matters. Accordingly, the following should be read as a tentative, preliminary exploration of the data that were available to us through our historiometric study of cases.

Table 5 presents the principal investigator's (DuBois) coding of these responses—treating responses to all three questions as generating one overarching "theory of the case."

Table 5: Qualitative Themes from the Research Misconduct Cases (n = 40)*

1. Self-centered personality or thinking	48%
a. Bright, arrogant, narcissistic	
b. Thought they would get away with it—it seemed easy	
c. Lust for scientific success—highly ambitious	
d. Greed—increase wealth (rather than just keep job with next grant)	
2. Unwarranted certainty—making data fit hypotheses that they thought were correct, making images more clearly represent 'the truth'	13%
3. Pressure—intense desire for a quick publication or next grant	33%
4. Careless—either with data, images, or supervision	25%
5. Unclear—uncertain about what they were thinking or what contributed to behavior	20%

*Percentages add up to more than 100% because a second code was used in some cases

DISCUSSION

We examined 40 high profile cases of research misconduct and compared them to 40 cases of other violations of research ethics (heavily representing human subjects violations) and 40 cases of wrongdoing in medical practice. To summarize our findings: Cases of research misconduct generally involved repeat offenses (68%) by an individual who was acting alone (90%) across an average of 3.8 years. In no cases did the individual plead insanity or blame the behavior on an addiction; only one case involved the claim to be following orders. Twenty-eight percent of cases involved failed initial attempts to report misconduct—that is, suspicions were reported, but either not investigated or no finding was made initially and the behavior continued. In contrast to other kinds of wrongdoing, institutional colleagues comprise the largest group of whistleblowers in misconduct cases. Only two environmental variables scored higher than 1.5 on a scale from 1 to 3: Oversight failures and being in a position of authority. Both of these factors—as well as financial rewards—were correlated with more serious cases of misconduct—but none of these factors was more strongly present in misconduct cases than other research wrongdoing cases.

When compared to other kinds of wrongdoing in research or to violations of medical ethics, research misconduct is associated with very few variables. This is somewhat surprising given the number of variables that this project tracked, and the fact that we explored some novel environmental factors—including "opportunity" factors—that might explain how research misconduct arises.

In what follows we discuss the implications of our findings for responsible conduct of research (RCR) education, for research integrity officer (RIO) behavior, and for further research.

RCR Education

By studying cases of research misconduct, we have learned important lessons for RCR education, lessons that would justify recent shifts from

philosophically to psychologically driven programs. The norms involved in misconduct cases were never ambiguous; this contrasted significantly with other kinds of wrongdoing in research. Accordingly, it is unlikely that traditional ethics education—with its focus on promulgating specific rules as well as identifying general principles and strategies for deducing specific rules—will help prevent or remediate misconduct.

Individual factors appear to play a significant role in research misconduct. The "history of wrongdoing" results did not suggest that misconduct is the product of antisocial personalities (e.g., only 3% had felony arrests). However, qualitative data suggest that narcissistic thinking plays a role: We repeatedly observed senior investigators who thought selfishly, thought they could get away with something, or thought they were justified in making data fit their hypotheses. We also found wrongdoers reporting pressure to publish and obtain grant funding; in some cases, however, this also tied into narcissistic thinking insofar as they did not aim merely to keep their jobs with publications and grants, but to become superstars in their fields. Finally, we found some cases that involved carelessness—either with data or oversight of personnel.

Given these factors, we recommend that education focus more on sense-making strategies and mental models or bias reduction. Sense-making strategies teach people to question their own judgments, examine their motives, think of others, forecast consequences, and seek help (Mumford et al., 2008). Bias reduction education often requires group work aimed at identifying the distortions in selfish thinking (Gibbs, 2009).

While we offer these recommendations for education addressing research misconduct, they seem sensible regarding all areas of the responsible conduct of research. However, more traditional ethics education may have a greater role to play in areas with ambiguous norms, for example, the negotiation of order of authorship or best practices for enrolling participants who lack decisional capacity.

RIO Behavior

Given that most whistleblowers are institutional colleagues or subordinates, RIOs should make themselves known and accessible to investigators. The fact that 28% of cases involved failed attempts at reporting suggests that individuals should be encouraged to report to those best qualified and most interested in following up on reports of suspicious data (that is, RIOs), rather than those most convenient to them within a research division. Additionally, our data suggest that cases of research misconduct are fairly heterogeneous. Given that most RIO's will investigate only a few cases of misconduct across a 5-year period, we recommend caution in generalizing about the causes and correlates of research misconduct.

Further Research

In some domains of wrongdoing, for example, wrongdoing in medical practice, we have identified strong environmental correlates to wrongdoing. Under such circumstances, it makes sense to complete a larger set of cases (e.g., 100 or more) to enable regression analyses and causal modeling. However, the failure to identify multiple moderate correlates to research misconduct suggests that further research with larger samples and regression modeling may yield little new information. We believe larger sample historiometric research is indicated to understand cases of human subjects protection failures and other kinds of wrongdoing in research. However, research on research misconduct might more fruitfully focus on individual variables such as cognitive biases, forecasting skills, and personality traits.

Policy and Prevention

We supported Adams and Pimple's (2006) suggestion to focus on environmental factors that provide opportunity for crime. However, our efforts to identify such correlates have proved far more fruitful in domains other than research misconduct. In part, this may be due to the fact that most investigators perceive themselves to have the opportunity to engage in research misconduct: They often enjoy high levels of autonomy, have access to databases, and decide on the final dataset that biostaticians and others see. But when such factors are ubiquitous they are also unlikely to emerge as statistical predictors. In this sense, some of the most prominent environmental factors that provide opportunity may be analogous to oxygen: You cannot have a fire without oxygen, but the presence of oxygen alone does not predict fire.

From a policy and prevention perspective, it may be that the key to prevention has more to do with increasing the likelihood of detection (that is, reducing the ubiquity of perceived opportunity) than with adjusting other features of the environment (such as financial rewards, social norms, or conflicting roles).

REFERENCES

Adams, D., and Pimple, K. (2005). Research misconduct and crime lessons from criminal science on preventing misconduct and promoting integrity. *Accountability in Research* 12: 225–240.

Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Pers.* Soc. Psychol. Rev. 3: 193–209.

Bramstedt, K. A., and Kassimatis, K. (2004). A study of warning letters issued to institutional review boards by the United States Food and Drug Administration. *Clinical and Investigative Medicine. Medecine Clinique et Experimentale* 27: 316–323.

Brewer, M. (2000). Research design and issues of validity. In *Handbook of Research Methods in Social and Personality Psychology*. Reis, H., and Judd, C. (eds.) pp. 3–16. Cambridge: Cambridge University Press.

Dana, J., and Loewenstein, G. (2003). A social science perspective on gifts to physicians from industry. *JAMA: Journal of the American Medical Association* no. 290: 252–255.

Davis, M. S. (2003). The role of culture in research misconduct. *Account*. Res. 10: 189–209.

Davis, M. S., Riske-Morris, M., and Diaz, S. R. (2007). Causal factors implicated in research misconduct: Evidence from ORI case files. *Sci. Eng. Ethics* 13: 395–414.

Depp, C., and Lebowitz, B. D. (2007). Clinical trials: Bridging the gap between efficacy and effectiveness. *International Review of Psychiatry* 19: 531–539.

DuBois, J. M., Anderson, E. E., Carroll, K., Gibb, T., Kraus, E., Rubbelke, T., and Vasher, M. (2012). Environmental factors contributing to wrongdoing in medicine: A criterion-based review of studies and cases. *Ethics Behav*. 22: 163–188.

DuBois, J. M., Kraus, E., and Vasher, M. (2012). The development of a taxonomy of wrongdoing in medical practice and research. *American Journal of Preventive Medicine* 42: 89–98.

DuBois, J. M., Anderson, E. E., and Chibnall, J. (2013a). Assessing the need for a research ethics remediation program. *Clinical and Translational Science* 6(3): 209–213.

DuBois, J. M., Anderson, E. E., and Chibnall, J. T. (2013b). Understanding the severity of wrongdoing in healthcare delivery and research: Lessons learned from a historiometric study of 100 cases. *AJOB Primary Research*.

Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS One* 4: e5738. doi: 10.1371/journal.pone.0005738.

Gibbs, J. C. (2009). *Moral Development and Reality: Beyond the Theories of Kohlberg and Hoffman*. New York: Pearson.

Greenberg, J. (1993). Stealing in the name of justice: Informational and interpersonal moderators of theft reactions to underpayment inequity. *Organizational Behavior and Human Decision Processes* 54: 81–103.

Grover, S. L. (1993). Why professionals lie: The impact of professional role conflict on reporting accuracy. *Organizational Behavior and Human Decision Processes* 55: 251–272.

Hegarty, W., and Sims, Jr., H. (1978). Some determinants of unethical decision behavior: An experiment. *Journal of Applied Psychology* 63: 451–457.

Jansen, E., and Von Glinow, M. A. (1985). Ethical ambivalence and organizational reward systems. *The Academy of Management Review* 10: 814–822.

Jennings, M. (2006). *The Seven Signs of Ethical Collapse: How to Spot Moral Meltdowns in Companies . . . before It's too Late*. New York, NY: St. Martins Press.

Katz, D., Caplan, A. L., and Merz, J. F. (2003). All gifts large and small: Toward an understanding of the ethics of pharmaceutical industry gift-giving. *Am. J. Bioeth*. 3: 39–46.

Keith-Spiegel, P., and Koocher, G. P. (2005). The IRB paradox: Could the protectors also encourage deceit? *Ethics Behav*. 15: 339–349.

Levine, R. J. (1992). Clinical trials and physicians as double agents. *Yale Journal of Biology and Medicine* 65: 65–74.

Marshall, E. (1999). Clinical research: Shutdown of research at Duke sends a message. *Science* 284: 1246–1246.

Martinson, B. C., Anderson, M. S., Crain, A. L., and De Vries, R. (2006). Scientists' perceptions of organizational justice and self-reported misbehaviors. *J. Empir. Res. Hum. Res. Ethics* 1: 51–66.

Martinson, B. C., Anderson, M. S., and de Vries, R. (2005). Scientists behaving badly. *Nature* 435: 737–738.

Martinson, B. C., Crain, A. L., De Vries, R., and Anderson, M. S. (2010). The importance of organizational justice in ensuring research integrity. *J. Empir. Res. Hum. Res. Ethics* 5: 67–83.

Meyer, W. M., and Bernier, Jr., G. M. (2002). Potential cultural factors in scientific misconduct allegations. In *Investigating Research Integrity: Proceedings of the First ORI Research Conference on Research Integrity*. Steneck, N. H., Scheetz, M. D. (eds.) pp. 163–166. Washington, D.C.: Department of Health and Human Services.

Mumford, M. D. (2006). *Pathways to Outstanding Leadership. A Comparative Analysis of Charasmatic, Ideological, and Pragmatic Leaders*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Mumford, M. D., Connelly, M. S., Helton, W. B., Strange, J. M., and Osburn, H. K. (2001). On the construct validity of integrity tests: Individual and situational factors as predictors of test performance. *International Journal of Selection and Assessment* 9: 240–257.

Mumford, M. D., Connelly, S., Brown, R. P., Murphy, S. T., Hill, J. H., Antes, A. L., Waples, E. P., and Devenport, L. D. (2008). A sensemaking approach to ethics training for scientists: Preliminary evidence of training effectiveness. *Ethics Behav*. 18: 315–339.

Mumford, M. D., Devenport, L. D., Brown, R. P., Connelly, S., Murphy, S. T., Hill, J. H., and Antes, A. L. (2006). Validation of ethical decision making measures: Evidence for a new set of measures. *Ethics Behav*. 16: 319–345.

Office of Statistical Analysis and Reporting. (2012). Research Grants: Awards, by Gender. In *NIH Databook*. Bethesda, MD: National Institutes of Health.

Rodwin, M. A. (1993). *Medicine Money and Morals: Physicians' Conflicts of Interest*. New York: Oxford University Press.

Schuchman, M. (2008). Medical whistle-blower protection lacking. *Canadian Medical Association Journal* (12): 1529.

Shah, S., Whittle, A., Wilfond, B., Gensler, G., and Wendler, D. (2004). How do institutional review boards apply the federal risk and benefit standards for pediatric research? *JAMA: Journal of the American Medical Association* 291: 476–482.

Silvia, P. J. (2005). What is interesting? Exploring the appraisal structure of interest. *Emotion* 5: 89–102.

Simonton, D. K. (2003). Qualitative and quantitative analyses of historical data. *Annual Review of Psychology* 54: 617–640.

Titus, S. L., Wells, J. A., and Rhoades, L. J. (2008). Repairing research integrity. *Nature* 453: 980–982.

Victor, B., Trevino, L. K., Shapiro, D. L. (1993). Peer reporting of unethical behavior: The influence of justice evaluations and social context factors. *Journal of Business Ethics* 12: 253–263.

Zimbardo, P. G. (2007). *The Lucifer Effect*. New York: Random House, Inc.