The Past, Present, and Future of Scientific Misconduct Research: What Has Been Done? What Needs to be Done?

Jungmin Lee
Vanderbilt University

Abstract: This literature review examines previous studies on scientific misconduct to answer the three following questions: 1) What is scientific misconduct and how prevalent is it in the U.S.? 2) How does the social control mechanism detect, deter, and sanction scientific misconduct? 3) At both individual and institutional levels, what hinders the social control mechanism? Due to the lack of empirical studies on scientific misconduct, this study explores survey results and reports from governmental agencies as well as scholarly papers. In addition to answering the three questions above, this review identifies limitations in the literature and provides suggestions for future research in this area.

Autonomy and academic freedom are considered as central features of the academic profession. These privileges were given and supported based on the assumption of self-regulation. In other words, the general public believes that the academic community is able and willing to regulate its own members’ misbehaviors (Braxton, 1999). Scientists also tend to believe that scientific misconduct is very rare and it can be self-regulated due to the competitive nature of the scientific community. Even if a scholar published highly innovative results based on data fraud, his misconduct would be revealed by other scholars who became skeptical and investigated the result (Steneck, 1994). However, this long-held belief was challenged after several cases of research misconduct that were revealed in 1980 (Dresser, 2001). Revelation of these cases raises a need for setting an explicit standard about scientific misconduct and regulations.

Jungmin Lee is a Ph.D. Student in the Higher Education Leadership and Policy Program at Vanderbilt University.

Copyright © 2011 by The Journal of the Professoriate, an affiliate of the Center for African American Research and Policy. All Rights Reserved (ISSN 1556-7699)
Since then, a variety of research regarding scientific misconduct has been implemented. However, most of the research focuses on the definition of scientific misconduct or policy recommendations based on literature reviews. This paper reviews previous literature on scientific misconduct and suggests what kind of research needs to be conducted for future research. This paper consists of three parts as follows: 1) scientific misconduct 2) social control mechanisms of scientific misconduct 3) factors that hinder social control of scientific misconduct. The first section explores the definition, the prevalence, the subject and possible explanation of scientific misconduct. The next section employs the framework of social control mechanism that is widely used in sociology: detection, deterrence, and sanction (Zuckerman, 1988). Using the framework, this review examines how this mechanism suppresses scientific misconduct. In reality, the social control mechanism appears not to work very well: a number of scientific misconduct cases are reported every year, and more seriously, even more cases might never be reported. The third section investigates what factors hinder the social control mechanism. In each section, this paper examines previous studies and their limitations. The author provides a suggestion for future research in the concluding section.

**Scientific Misconduct**

**Definition of Scientific Misconduct**

Most research about scientific misconduct focuses on the definition of scientific misconduct. The major argument here is whether or not to extend the definition beyond falsification, fabrication, and plagiarism (FFP). While the National Science Foundation (NSF) and the Office of Research Integrity (ORI) extended the definition to “other practices that are seriously deviated from accepted”, the National Academy of Sciences (NAS) insisted to limit the scientific misconduct to FFP (Decoo, 2002).

Rationales from each side are clear. Buzzelli (1993) argues for the open-ended phrase not to leave room for protection for a potential perpetrator. As he mentioned, it is impossible to predict every possible case of scientific misconduct in advance. Thus, it is necessary to include the phrase to regulate scientific misconduct beyond FFP. In contrast, those who favor the limited version warn that the inclusive version would prevent scientists from employing “unorthodox, highly innovative approaches that lead to major advances in science” (Schachman, 1993).
They admit that other wrongdoings such as sexual harassment of students can happen, and they should be punished. However, they should not be defined as “misconduct in science” if they are “not unique to the conduct of science” and can be regulated by other policies (Schachman, 1993).

Although there have been fierce discussions based on the rationales, some scholars consider these arguments to be about a power game between the academic community and the government rather than right or wrong (Steneck, 1994). As mentioned, academic professionals value their autonomy and academic freedom. However, when the self-regulation system didn’t work out properly as in 1980s, they had no choice but to accept supervision and regulation from a third party such as federal government agencies (Steneck, 1994). Despite this, they tried to limit the level of regulation to be as low as possible. If they allowed the open-ended phrases, it may lead to additional regulations of government in the future. This might be one reason why it took so long to agree on the unified definition.

Finally, the battle ended in favor of the academic community. In 2000, the Office of Science and Technology Policy announced the definition of scientific misconduct, which is limited to “FFP in proposing, performing, or reviewing research, or in reporting research results”. In addition, any misconduct committed without intent is not considered as scientific misconduct. This definition is applied to any research supported by U.S federal agencies (Decoo, 2002).

In spite of the unified federal definition, problems regarding the definition of scientific misconduct still remain. First, the new definition does not cover the whole range of misconduct. For example, ORI surveyed principal investigators who are funded by National Institutes of Health (NIH) about scientific misconduct that they have observed. About 24% of the scientific misconduct cases reported did not fall under the FFP definition. Rather, they were questionable research practices affecting data, plagiarism of assignments or slides, IRB and human subject issues, conflict of interest issues, or authorship issues. Although the author named these as “research misbehavior”, this example implies the gap between the federal definition and the perception of researchers in fields (Wells, 2008).
Second, the definition of scientific misconduct also varies at the institutional level. The ORI reports that other unacceptable research practices are still included as scientific misconduct in several institutional policies. For example, 33% of institutional policies reviewed define the failure to comply with governmental regulation as scientific misconduct. Improper use of authorship\(^1\) (15%) and retaliation to whistleblowers (16%) are also considered as misconduct in science in some institutions (Center for Health Policy Studies, 2000). This gap in the definition between federal regulation and institutional policy may confuse researchers and students in institutions. The confusion may hinder them from making allegations, and thus, fail to detect misconduct.

**Prevalence of Scientific Misconduct**

There are two perspectives regarding the prevalence of scientific misconduct: rotten apple and the tip of iceberg (Anderson, 1999). As mentioned earlier, most scientists tend to believe the rotten apple view. However, we cannot judge it because the prevalence of scientific misconduct is not accurate either. As Hackett (1999) mentions, there are two major sources of statistics regarding scientific misconduct. The first one is the report of federal agencies such as the ORI and the OIG (Office of Inspector General), and the other one is survey results based on self-report. However, none of these statistics can be accurate enough to capture the exact number of incidents of scientific misconduct.

Regarding the data from federal government agencies, each case reported is clearly scientific misconduct. They report the number of allegations and misconduct cases based on what they receive from individual institutions and whistle blowers. If they announce a case as misconduct, it went through the inquiry and investigation process based on strong evidence. However, the number of incidents of misconduct in general is underreported (Altman, 1997): Principal investigators who responded to the ORI survey reported that 36% of the misconduct that they observed was not reported (Wells, 2008). In addition, data of the oversight agencies highly depend upon institutional report. However, an institution

---

\(^1\) It is defined as “improper assignment of credit such as excluding authors, misrepresentation of the same material as original in more than one publication, inclusion of individuals as authors who have not made a definite contribution to the work published; or submission of multi-authored publications without the concurrence of all authors” (Center for Health Policy Studies, 2000).
does not need to report misconduct if it decides the case is not worth opening an investigation. In other words, there could be many suspected cases that are silently handled within an institution.

Statistics from self-reporting surveys are not reliable, either. Due to the social desirability issue, most questionnaires ask researchers if they “observe or have direct knowledge” of other’s misconduct. Since some of the reported misconduct might not be investigated, we cannot be sure if their suspected misconduct is actually scientific misconduct. That is, some of the reported misconduct in a survey might be based on problematic information. Moreover, the number of cases can be inflated. Suppose that one case of scientific misconduct was revealed in a department. If two professors in the department are asked to count the number of acts of scientific misconduct that they have observed, each of them might report the same case. In this case, a researcher who collects the questionnaire will count them as two cases because he does not know that they refer to the same case. Lastly, low response rates of surveys regarding scientific misconduct raise a question about the generalizability of its result (Wells, 2008).

Due to the limitations of each source, the only conclusion we can draw from data is the fact that scientific misconduct is not as rare as most scientists believe. Altman (1997) introduces results of several large-scale surveys about scientific misconduct, and all of them reported a substantial amount of scientific misconduct. For example, the “Project on Professional Values and Ethical Issues in Graduate Education of Scientists and Engineers” surveyed 2,000 doctoral candidates and 2,000 professors awarded federal grants. Among them, 6% of doctoral students and 9% of faculty reported that “they have direct knowledge of faculty members involved in data falsification or plagiarizing” (Altman, 1997). Additionally, 7.4% of principal investigators funded by NIH also reported that they observed at least one case of scientific misconduct during the last three years (Wells, 2008). Graduate students are also exposed to a number of misconduct in school: According to Anderson, Louis, and Earle (1994), on average, graduate students surveyed have seen two to five misconduct cases committed by their peer students or faculty members. These results suggest the need to pay attention to scientific misconduct rather than considering it as a rare accident.
Subject and Explanations for Scientific Misconduct

Due to the low reliability of existing data, it is very difficult to find a common characteristic of misconduct perpetrators or the main cause of committing misconduct. According to ORI data, there is no specific characteristic of perpetrators. Based on reported cases, perpetrators can be found in most academic ranks or tenure status (Rhoades, 2004). However, it is necessary to keep in mind that this information is only suggestive.

Zuckerman (1988) provides four different explanations for scientific misconduct: anomie theory, psychopathology, conflict theory, and other structural sources. Anomie theory interprets scientific misconduct as an attempt to gain recognition by using “illicit means”. For example, perpetrators highly value peer recognition, but they are unable to achieve it with valid means, which leads to their engagement in scientific misconduct. Psychopathology attributes scientific misconduct to the “pathological self-deception” of an individual scientist. Because these scholars are firmly convinced of their conclusion, it is not a big problem for them to manipulate or invent evidence that support their conclusion. Conflict theory asserts that scientists are alienated from their industrialized work place and became involved in misconduct when the regulation is loose. Structural sources include the pressure for research grants and publications, the lack of time to check others’ work (e.g., graduate students, post-doc, or co-authors), and the nature of interdisciplinary research such that researchers from different fields have difficulty in cross-checking.

To date, most empirical studies found the structural sources explain scientific misconduct. Mostly, a researcher’s high level of stress or pressure might lead to scientific misconduct. Green and Goldberg (1994) reported that researchers who work for profit organizations are more likely to observe scientific misconduct than their counterparts who work in the state government. After reviewing scientific misconduct cases that were committed by research trainees, Wright, Titus, and Cornelison (2008) concluded that most of the trainees were under high levels of internal pressure to publish a paper or show their achievement. Also, they reported that their mentors rarely examine the raw data or that there was a lack of standard regarding research ethics in their lab. This suggests that high levels of stress coupled with the lack of appropriate mentoring contributes to scientific misconduct.
Above all, the very limited access to perpetrators of misconduct hinders further research in this area. Because it is difficult to contact perpetrators, most of the studies in this area depend on ORI documents or observers of misconduct. Thus, at best, their results can be interpreted as correlation rather than a causal relationship. Moreover, exploring the context of observers might not help in understanding the cause of scientific misconduct because committing scientific misconduct is more of an individual matter (Anderson, et al., 1994).

**Social Control Mechanism of Scientific Misconduct**

**Detection & Deterrence**

Detecting scientific misconduct is a direct means to control deviant behaviors. It also works as a deterrence mechanism if other researchers are aware of the possibility to be detected. The most effective way of detecting misconduct is replication. Scientific misconduct can be suspected if a research result is not reproducible. However, due to the high emphasis on originality, scientists more often find suspicious evidence accidently by extending other scientists’ research rather than by replicating the previous research directly. Thus, the more significant and innovative a research is, the more it is likely to be replicated. Although it is difficult to prove the deceitful intention of a possible perpetrator from the “inadvertent replication”, raising a question on the result is enough to protect the knowledge in the field. The possibility that any results can be replicated deters scientists from committing misconduct (Zuckerman, 1988). For these reasons, Fox and Braxton (1999) recommended rewarding the actual replications by scientists. In addition, peer reviews can be used to detect researcher’s failure to give credit to other scholars although it is not effective to detect other types of research misconduct (Fox & Braxton, 1999; Zuckerman, 1988).

Since it is difficult and costly to detect and report scientific misconduct, the best way to decrease misconduct is preventing it in advance rather than regulating it afterwards (Fox & Braxton, 1999). Education and training are critical factors in preventing future misconduct. In particular, many researchers emphasize the importance of socialization in graduate school (Folse, 1991; Fox & Braxton, 1999; Braxton & Baird, 2001). As Anderson et al. (1994) says, it is graduate school where “students learn, formally and informally, what behaviors are expected and rewarded in academic research and what constitutes unacceptable deviation from
shared norms of conduct” (p.331). Although Anderson et al. (1994)’s empirical analysis failed to find empirical evidence, Victor and Cullen (as cited in Anderson et al., 1994) find that the ethical climate of an organization affects ethical behaviors of its members.

Institutional policy can be another mechanism of deterrence and detection: being well aware of institutional policy prevents researchers from involving themselves in misconduct and encourages them to report suspected misconduct. Wells (2008) supported this view by reporting that scientists who read their institutional policies are more likely to make allegations. Since the 1985 Health Research Extension Act, which required universities to develop their own institutional policy to deal with scientific misconduct, almost every institution funded by the federal government has its own policy. Moreover, 78.4% of accredited medical schools in the U.S have their own guideline (Douglas-Vidas, Ferraro, & Reichman, 2001). These figures imply that establishing an institutional policy regarding research ethics is stabilized at the institutional level. The problem is the policy’s content and usage.

One ORI report in 2000 reviewed 156 institutional policies consistent with the ORI guideline and found that only 29% of them explicitly state the duty of reporting scientific misconduct. Also, except for two institutions, no policy clearly states if the institution will pursue anonymous allegations (Center for Health Policy Studies, 2000). These ambiguous statements in institutional policies might hinder the attempts of potential whistleblowers to make allegations. Also, Douglas-Vidas et al. (2001) found that 32% of medical school guidelines only mention legal issues such as intellectual property. This result suggests that the guidelines in medical schools were developed to deal with legal conflicts rather than to prevent research misconduct.

Sanction

Zuckerman (1988) mentions that mechanisms for sanction are less developed than the mechanisms for deterrence and detection. As discussed above, formal regulations on research misconduct are recently established and often unclear about critical matters. Moreover, there is no consensus about how far a principal investigator should be responsible for research misconduct committed by his coauthors or students. Due to the ambiguity, the level of sanction on the same kind of misconduct might vary across individuals, departments, or institutions.
For instance, a department chair can choose the level of sanction toward a perpetrator in his department. Several studies explore what factors decide formal or informal action of a professor (Braxton, 1991; Braxton & Bayer, 1994; Braxton & Bayer, 1996). Formal action means more open and severe sanction such as a formally reported file or terminating the perpetrator’s employment. Results suggest that professors who are more likely to take formal actions are those with more publications, a high level of professional status, a position in a department chair, or are in charge of a department chair in a renowned department. In contrast, professors who value symbolic criteria or professional solidarity, who are afraid of stigmatization as a whistleblower, or who graduate from a prestigious department tend to take informal action (Braxton, 1991; Braxton & Bayer, 1994; Braxton & Bayer, 1996).

Except the Braxton & Bayer’s study (1996) that asked about the formality of actions a professor actually took, most studies regarding sanction merely ask the hypothetical action of survey respondents: How would you respond in a specific situation? Although it is the best approach given the difficulty to contact people who were actually involved in scientific misconduct, results from these studies might not reflect the actual response of people. In other words, people can behave differently from what they say they would do.

**What Hinders Social Control of Scientific Misconduct?**

If the social control mechanism works, researchers or institutions that detect scientific misconduct should report it. However, in many cases, they do not. This section examines why people and institutions are reluctant to report or sanction scientific misconduct even after its allegation.

**Individual Level**

Braxton and Bayer (1994) provide several reasons why people do not report scientific misconduct. First, the definition of misconduct is confusing. If they cannot decide whether a certain action is misconduct or not, they would not report it. Second, people are worried about the negative consequences of reporting on their own life. According to the
nationwide survey\(^2\), 53% of students and 26% of faculty feel that they would expect retaliation after whistle blowing. The degree of anxiety was stronger if the respondent was 1) a student rather than a professor 2) a junior professor rather than a senior professor 3) making allegations against colleagues (other professors) rather than students. Lubalin, Ardini, and Matheson (1995) also found that 69% of whistleblowers experienced negative consequences after their reporting: It includes being pressured to drop or counter their allegations (42.6%/39.7%), being ostracized by their colleagues (25%), experiencing material deficit such as a decrease in research support (20.6%)/staff support (10.3%) or denial of salary increase (11.8%)/promotion (7.4%), and most seriously being fired or not renewed (11.8%). These negative consequences prevent people from making allegations: Braxton and Bayer (1996) found that faculty members who are afraid of being labeled as a whistleblower tend to take less formal actions\(^3\) toward research misconduct.

Third, a researcher who values a professional etiquette, a belief that colleagues should not criticize or hurt another’s reputation, would not report their colleague’s misconduct. Fourth, some scholars believe that any means can be justified for advance in science. Fifth, the level of action (sanction) is sometimes decided based on the functional and symbolic criteria. Functional criteria regard the sanction’s effect to prevent future misconduct and symbolic criteria consider if a sanction openly humiliates the perpetrator. If an observer thinks of these criteria, they would decide not to report what they have seen (Braxton & Bayer, 1996). Lastly, who holds the responsibility does matter. If a researcher were explicitly responsible for reporting misconduct, he would be more likely to report it. Although this assumption has not proven empirically in terms of scientific misconduct, other experiments about academic misconduct support the assumption. For example, Trevino and Victor (1992) explored a situation in which students would report their

---

\(^2\) The Project on professional values and ethical issues in graduate education of scientists and engineers

\(^3\) The authors measured the degree of formality of actions taken by faculty members who knew of other’s misconduct as follows: “none (=1), professionally boycott (=2), reported to other colleagues (=3), individual personally confronted (=4), formal report filed (=5), individual’s appointment terminated (=6)” (Braxton & Bayer, 1996).
classmates’ cheating. Results suggest that students are more willing to report the cheating if they are well aware of the honor code. Rennie and Crosby (2001) also found that medical school students would not report their colleagues’ misconduct because 1) it is not a student’s responsibility, 2) there is no relevant guideline, and 3) reporting is not acceptable in terms of their peer’s norm.

Institutional Level

Why is an institution reluctant to report misconduct of its constituents? First, it will damage the institution’s reputation (Altman, 1997). Second, it would take away the material resources of the institution. More specifically, an institution would lose its overhead cost or it should pay for penalty under the False Claim Act. Turner (1999) explains the meaning of overhead to university administrators. Once a researcher is awarded an external grant, a substantial portion of it flows into the university budget in the name of overhead cost. Of course, actual maintenance cost spent by the research project team (e.g., building maintenance, water and electricity usage) is much smaller than the amount of the overhead: A substantial amount of overhead cost can be used for other purposes of the university. Moreover, unlike endowment or gifts from alumni, overhead cost is not restricted to a specific use. This free-to-use nature of overhead gives a discretionary power to administrators over boards and alumni. Therefore, it is natural for them to avoid any chance of losing the money by reporting a recipient of an external grant.

The False Claim Act might have a more direct effect on an institution. According to the law, “Any person who knowingly presents a false or fraudulent claim for payment to an officer or employee of the U.S government … is liable to the U.S for up to three times the amount of monetary damages that the government sustains because of the false claim, plus a civil penalty of $5,000 to $10,000 for each false claim” (Sherman, 1995). In 1995, under this law, the University of Alabama in Birmingham was fined about two million dollars for its professor’s plagiarism in a grant application for NIH funds. Considering the legal responsibility of an institution and the astronomical figures, institutions would like to deal with any relevant misconduct silently (Sherman, 1995).
Conclusion & Implications for Future Research

This paper reviews previous studies on scientific misconduct and addresses three questions as follows: 1) What is scientific misconduct? 2) How does the social control mechanism detect, deter, and sanction scientific misconduct? 3) What hinders the social control mechanism? Since scientific misconduct was a controversial issue in the 1980s, a number of scholars and practitioners have attempted to investigate, analyze, and finally deter it. However, their attempts might not be feasible in the first place, not only because evidence of scientific misconduct is rarely conspicuous but also because the cost for finding others’ misconduct far exceeds any benefits that it brings to individual researchers.

Except for plagiarism, even experts cannot tell if a research result in their field is based on falsification or fabrication unless they find substantive evidence of misconduct by replicating the research. In addition, there is not much incentive for most scholars to replicate other scholars’ research, especially in today’s competitive academic environment. Given the current reward structure, replication itself is not considered as important as original scholarly work. Moreover, even if scholars found suspicious evidence of misconduct, they should be very cautious in making allegations. If the allegation turned out to be incorrect, they might feel guilty about their hasty judgment toward their colleagues. In the worst case, they could be accused of defamation and face substantial consequences of their wrong allegations.

Given the context, the role of whistleblowers is highly critical since most of them know of evidence of scientific misconduct. Unfortunately, the cost for internal whistle blowing is much higher than it is for outside experts who make an allegation. Whether or not their allegations turned out to be true, most whistleblowers were stigmatized by their colleagues, at best. In many cases, they experience several economic and career disadvantages due to their actions. In contrast, they, as individual scholars, benefit almost nothing from whistle blowing. The only benefit that they might have is a higher self-esteem or better mental health, which is too small to warrant the cost.

Therefore, unless the current peer-review system and the reward structure in academia are changed, it is not realistic to expect to
effectively deter further scientific misconduct. However, this conclusion does not mean that we should stop making efforts to deter misconduct and just let any researchers involved get away with this practice. Instead, scholars need to share and analyze previous misconduct cases in more detail in order to understand the mechanism of scientific misconduct. It is impossible to prevent, detect, and sanction every single misconduct case; however, we can improve the current system by giving an incentive for detecting misconduct, protecting whistleblowers or outside scholars who make allegations, and developing more effective sanctions that have a significant and long-lasting consequence for perpetrators. We cannot eliminate every single misconduct attempt, but we can significantly reduce them. Moreover, as long as scholars have autonomy over their work, it is their responsibility to control their colleagues’ faulty behaviors.

Given this responsibility, scholars need to illuminate a number of covert aspects of scientific misconduct. Among them, this review suggests implications for future research based on Anderson’s (1999) recommendation. Anderson (1999) provides four elements of scientific misconduct (context, misconduct, exposure, and consequences) and suggests directions for future research within a framework of the four elements. At first, Anderson (1999) recommends several directions for each element of scientific misconduct. For the misconduct element, it is necessary to extend the scope of misconduct research to errors and minor hypocrisies. Research regarding context needs to investigate contextual effect: In each misconduct case, which contextual factors are actually relevant to misconduct? For the element of exposure, two directions are suggested. First, exposure should be viewed as a process to further misconduct rather than a consequence of misconduct. In fact, perpetrators might produce other research fraud to conceal their prior misconduct. Second, considering the fact that most misconduct cases are handled informally even after being exposed, future research needs to investigate the “quiet exposure and correction of misconduct” (Anderson, 1999). Consequence elements need to be investigated along with the context of misconduct case and the way how the case is handled.

Although it has been ten years since Anderson suggest the future directions, only a few studies have been conducted as recommended. For example, studies of Green and Goldberg (1994) and Wright et al. (2008) attempt to find the context in which scientists are more likely to commit
scientific misconduct. Except this strand of research, Anderson’s other suggestions are not reflected in misconduct research to date.

Anderson (1999) also recommends the extension of the four elements: interactive effects, longitudinal effects, and broader contextual effects. Research about interactive effects needs to examine the four elements in each misconduct case and their interactive relations. Longitudinal effects focus on the subsequent effects of one misconduct case to latter misconduct cases: How does the consequence of a previous misconduct case affect the context of latter misconduct case? It is also recommended to keep track of a perpetrator’s career path before and after his misconduct. Broader contextual effects reflect the highly complicated research environment, such as a number of research regulations from external agencies, collaborated research that usually involves a number of co-researchers, and the normative structure of research. Exploring how these research environments affect the four elements of scientific misconduct (and their interactions) is needed.

To date, almost no study empirically investigates the interactive and longitudinal effects. As an individual researcher, it is almost impossible to have access to the whole process of a scientific misconduct case. Given the limited access, it is worth exploring a mechanism of scientific misconduct from a well-developed experiment that hypothesizes a specific context and asks respondents what they would do in that situation. Regarding the broader contextual effects, several studies have been conducted to examine the effect of normative structures on the level of sanction toward misconduct (Braxton & Bayer, 1994; Braxton & Bayer, 1996). However, as research environment is increasingly competitive, investigating its effects on scientists’ incentive to involve in misconduct would be worthwhile.
References


