東京大学 公共政策大学院
ワーキング・ペーパーシリーズ
GraSPP Working Paper Series
The University of Tokyo

GraSPP-DP-E-14-004 and STIG-DP-E-14-004

Dishonest Conformity In Peer Review

Sotaro Shibayama

Yasunori Baba

December 2014

GraSPP
THE UNIVERSITY OF TOKYO
GraSPP Discussion Paper E-14-004

STIG
SCIENCE, TECHNOLOGY, AND INNOVATION GOVERNANCE
STIG Discussion Paper E-14-004

GRADUATE SCHOOL OF PUBLIC POLICY
THE UNIVERSITY OF TOKYO
HONGO, BUNKYO-KU, JAPAN

GraSPP
THE UNIVERSITY OF TOKYO

GraSPP-DP-E-14-004 and STIG-DP-E-14-004

Dishonest Conformity In Peer Review

Sotaro Shibayama[a,b] , Yasunori Baba[a]

December 2014

[a] The University of Tokyo. Research Center for Advanced Science and Technology.
4-6-1 Komaba, Meguro-ku, Tokyo 153-8904, Japan
Phone/Fax: +81-3-5452-5371
e-mail: shibayama@00.alumni.u-tokyo.ac.jp
[b] The University of Tokyo.Department of Technology Management for Innovation.
7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan
Phone/Fax: +81-3-5452-5375
e-mail: baba@zzz.rcast.u-tokyo.ac.jp

GraSPP Discussion Papers can be downloaded without charge from:
http://www.pp.u-tokyo.ac.jp/research/dp/index.htm

Graduate School of Public Policy
The University of Tokyo
7-3-1, Hongo, Bunkyo-ku, Tokyo, 113-0033, Japan
Phone:+81-3-5841-1349

DISHONEST CONFORMITY IN PEER REVIEW

Sotaro SHIBAYAMA [a,b,*]

Yasunori BABA [a]

[a] The University of Tokyo. Research Center for Advanced Science and Technology.
4-6-1 Komaba, Meguro-ku,Tokyo 153-8904, Japan
Tel/Fax: +81-3-5452-5371
shibayama@00.alumni.u-tokyo.ac.jp

[b] The University of Tokyo. Department of Technology Management for Innovation.
7-3-1 Hongo, Bunky-ku,Tokyo 113-8656, Japan
Tel/Fax: +81-3-5452-5375
baba@zzz.rcast.u-tokyo.ac.jp

Abstract

Scientific honesty in scientific publication is critical for scientific advancement, but dishonesty is commonly and increasingly observed in the forms of misconduct and other questionable practices. Focusing on *dishonest conformity in peer review*, in which authors reluctantly obey referees' instructions in order to have their papers accepted even if the instructions contradict the author's scientific belief, this study aims to illustrate the nature of this specific form of dishonesty and examine the determinants of dishonesty. Drawing on survey data of Japanese life scientists, this study shows that the conflict between authors and referees in peer review is quite common and that a majority of scientists decide to follow referees' instructions rather than to rebut them. The results suggest that dishonest conformity occurs more often in basic biology than in medicine and agricultural sciences, when the corresponding author is under stronger scientific competition, if the author is an associate professor rather than a full professor, if the author has foreign research experience, and when the paper is submitted to low-impact journals rather than to medium-impact journals.

Keywords

Misconduct; publication; research integrity; scientific honesty; questionable research practice

---

[*] Corresponding author. This work was supported by the Japan Society for the Promotion of Science under Grant-in-Aid for Encouragement of Young Scientists (B) #24710160; and the Inamori Foundation. We acknowledge Seongwuk Moon and Scott Stern for their critical review for the earlier version of this paper.

1. Introduction

The progress of science is underpinned by the publication of scientific papers, whereby new discoveries are disseminated among members of the scientific community (David, 1998; Merton, 1973). As described as ‘standing on the shoulders of giants’ (Hess, 1997), science advances in a cumulative fashion based on the findings of previous research, which are shared through publication. Thus, it is critical, on the one hand, that reader scientists carefully scrutinize published findings (Merton, 1973), while on the other, that author scientists follow their scientific honesty (Zuckerman, 1977), the latter of which is the main subject of this essay.

Scientific honesty forms the core of research integrity. For example, the Singapore Statement on Research Integrity maintains ‘honesty in all aspects of research’ as one of the four principles,[2] and the National Institute of Health (NIH) grant policy refers to ‘honesty’ as a shared value.[3] Nevertheless, breaches of scientific honesty are not uncommon, ranging from apparent misconduct, such as fabrication and plagiarism, to other questionable practices, such as redundant publication and authorship abuse (Martin, 2013; Martinson et al., 2005). Fanelli (2009) estimates that approximately 2% of scientists have at least once committed fabrication, falsification, or modification of data or results. Fang et al. (2012) show that the percentage of retracted publications due to suspected fraud has increased by 10 times from the 1970s to the 2010s in biomedical and life sciences.

Dishonest publications themselves compromise the scientific knowledge base, and they also exacerbate the problem by inviting subsequent studies on an invalid basis. Studies based on false prior findings lack a logical foundation at least partially, or worse, can turn out false altogether,

[2] http://www.singaporestatement.org/statement.html accessed on August 25 2014.
[3] http://grants1.nih.gov/grants/research_integrity/whatis.htm accessed on August 25 2014.

which is a serious waste for the scientific community. Evaluating this negative impact of dishonest publications, Azoulay et al. (2012) show that after a paper was retracted for misconduct, other papers that had cited the retracted paper suffered a 5-10% decrease in the rate of citations they received. Preventing such grave consequences requires understanding why scientists deviate from the norm of honesty and taking necessary actions. However, the determinants of dishonesty are understudied though intense competition and pressure for publication are typically considered likely causes (Anderson et al., 2007b; Anderson et al., 2007c; De Vries et al., 2006). Despite attempts at some preventive measures, such as journal ethics policies and responsible conduct of research (RCR) training (Resnik and Master, 2013), their effectiveness has been questioned (Anderson et al., 2007a).

Thus far, research on scientific dishonesty has had two major limitations; that is, observed incidents of dishonesty are still rare, though increasing, and measurements of dishonesty suffer from desirability bias (Van Noorden, 2011). These are particularly formidable for the 'big three' of misconduct: fabrication, falsification, and plagiarism, on which the prior literature has primarily focused. In this regard, so-called questionable research practices are considered only minor deviations and are by far more common (Broad, 1981; De Vries et al., 2006; John et al., 2012; Martin, 2013; Martinson et al., 2005; NAS, 1992), and thus, we suppose that they could offer us an avenue to advance our understanding of scientific dishonesty. This study focuses on, among the many forms of questionable practices, dishonest conformity in peer review, in which authors reluctantly obey revision instructions in order to have their papers accepted even if doing so contradict the author's scientific belief.

Peer review is a critical device for maintaining the integrity of scientific publication, but a number of scientists believe that the current peer review system needs improving (Ellison, 2002a, 2002b). Frequently heard complaints that the review process is slow or that judgments are biased judgment are attributed to the misalignment of incentives for editors and especially for referees (Pitsoulis and Schnellenbach, 2012), but authors also act strategically and may behave dishonestly, as is the case in dishonest conformity. Though dishonest conformity in peer review has long been known (Frey, 2003) and anecdotally seems prevalent, empirical evidence is scant. Thus, this study aims to empirically investigate the practice of dishonest conformity to add to the literature on research integrity and on the peer review system.

2. Research integrity and dishonesty

Scientific honesty plays a crucial role in the science system, for the fundamental building blocks of the science system, including journal peer review and other evaluation systems, rely on the assumption that scientists follow their intellectual honesty (Zuckerman, 1977). Scientists in a position to evaluate publications do not usually bother confirming that the authors are not lying because detecting and proving dishonesty is prohibitively costly. Although the latest technologies such as software for detecting image manipulation may help find out a trace of dishonesty (Van Noorden, 2011), carefully worked-out manipulations cannot be easily identified, and importantly, policing such wrongdoings is not considered the function of peer review. Instead of implementing an elaborate policing system, the science system has been relying on heavy sanctions against wrongdoers. That is, culprits of misconduct, if caught, may be ostracized from the scientific community, and they can be prosecuted by funding agencies. These sanctions are formalized in some countries but remain informal in other countries. Some of the former countries have established special organizations for investigating misconduct, such as the Office of Research Integrity (ORI) in the US, the UK Research Integrity Office, and the Danish Committee on Scientific Dishonesty (Resnik and Master, 2013). What has been increasingly emphasized is preventive measures carried out by professional associations, academic journals, and funding agencies, which articulate research ethics codes and policies (Bosch et al., 2012; Macrina, 2007), and by research organizations, which provide education programs for RCR (Resnik and Master, 2013).

Nevertheless, reported incidents of misconduct have been increasing. Using bibliometric data of retractions, Fang et al. (2012) show that about two-thirds of retractions are attributable to misconduct and that the frequency of misconduct has increased by a factor of 10 since 1975 in the fields of biomedicine and life sciences. Similarly, Grieneisen and Zhang (2012), drawing on retraction data in various fields, suggest that 47% of retractions are caused by publishing misconduct (e.g., plagiarism, authorship issues), 20% by research misconduct (e.g., falsification, fabrication), and 42% by questionable data or interpretation and that the number of retractions increased by a factor of 20 from 2001 to 2010.

Prior literature and policy actions on dishonesty have focused primarily on fabrication, falsification, and plagiarism because of their relatively clear definitions.[4] However, other forms of questionable practices such as fragmentation of publication, redundant publication, and authorship abuse have long been noticed (Broad, 1981; Martin, 2013; Marusic et al., 2011). As early as 1992, the National Academy of Sciences (NAS) called attention to questionable practices that may not meet the definition of misconduct but that can erode confidence in research integrity, violate scientific traditions, affect scientific conclusions, waste time and

[4] For example, the definition of misconduct by National Science Foundation (http://www.nsf.gov/oig/resmisreg.pdf accessed on Aug 25 2014) and by NIH (http://grants.nih.gov/grants/research_integrity/research_misconduct.htm accessed on Aug 25 2014) lists only the big-three.

resources, and weaken the education of scientists (NAS, 1992). Compared to the big-three of misconduct, these practices are considerably more common (Martinson et al., 2005) and probably detectable with limited bias even by self-reporting surveys since they are not strictly considered misconduct. Taking advantage of these features, Martinson et al. (2005) identified several questionable practices and estimated their frequencies; for example, 27.5% of NIH-funded scientists self-reportedly engaged in 'inadequate record keeping related to research projects,' 15.3% 'dropped observations or data points from analyses based on a gut feeling that they were inaccurate,' and 7.6% 'circumvented certain minor aspects of human-subject requirements' at least once within three years. Further, a line of research has attempted to elucidate the mechanisms behind dishonesty. For example, competition (De Vries et al., 2006), seniority (Martinson et al., 2005), and perceived inequality in resource allocation (Martinson et al., 2006) are suggested as being responsible for some forms of questionable practices. Still, the literature on misconduct and other questionable practices is mostly conceptual or descriptive.

3. Dishonesty in Peer Review

Dishonesty can occur in various phases of research. Two of the big-three, falsification and fabrication, occur in producing and analyzing data, while plagiarism occurs in writing or publication. Authorship issues, fragmentation, redundant publication, and self-plagiarism also occur in publication. Our focal practice, dishonest conformity, occurs in the process of peer review.

Peer review is supposed to support research integrity, in which submitted papers are carefully scrutinized, necessary revisions are suggested, and acceptance or rejection for publication is recommended by expert referees. Though this system has been used for a long time, scientists have not always been content with it (Ellison, 2002a, 2002b). Peer review is based on volunteer referees but their incentives are often misaligned (Pitsoulis and Schnellenbach, 2012). Referees have little reason to provide accurate and timely reports, which delays the review process (Ellison, 2002b) and results in erroneous decisions (Coupe, 2004). Referees may be tempted to steal the results of submitted papers (Hagstrom, 1974). Taking advantage of their veto power, referees may make dishonest instructions. For example, they may write a negative report for rivals' papers (Ellison, 2002b) and require revisions favorable to their own theories. Relatively speaking, editors' incentives may be better aligned (Frey, 2003), but they also contribute to a similar set of problems. For example, Wilhite and Fong (2012) find that journal editors in the social sciences often coerce authors to cite the editors' papers to improve their own citation score. Since editors' reputations can improve by raising the status of their journals, they are incentivized to accept papers with a high probability of receiving citations (Frey, 2003), which can cause a bias in publishable findings. It is well known that editors are unwilling to accept low-impact results such as negative data and replication studies (Csada et al., 1996; Fanelli, 2010). In an attempt to address these problems, alternative publication systems have been proposed (Prufer and Zetland, 2010), and some of them have been actually implemented, such as post-publication open peer review (Ietto-Gillies, 2012). Nevertheless, the traditional peer review system remains the norm.

Knowing the problematic incentive structure, authors may also resort to questionable practices. For example, some journals are known to reject articles with imperfectly consistent results, which discourages authors from reporting inconvenient results. In addition, authors may reluctantly revise the content of their paper in order to have it accepted even if doing so contradicts the author's scientific belief. Describing such a practice, Frey (2003) coined the term 'academic prostitution.' Provided that most peer review journals bestow veto power to referees or editors, Frey (2003) contends that 'authors only get their papers accepted if they intellectually prostitute themselves by slavishly following the demands made by anonymous referees.' This is not to say that referees are always exploitative, but we anecdotally perceive that authors often encounter a dilemma of choosing whether to unwillingly follow referees or to risk rejection by not following referees (Frey, 2003; Tsang and Frey, 2007). In fact, the above-mentioned coercive citation, when cited papers are scientifically irrelevant, is a form of dishonest conformity. Nevertheless, dishonest conformity has not been investigated empirically, to the best of our knowledge, apart from the case of coercive citations (Wilhite and Fong, 2012).

Conformity in the above setting may not be equated to dishonesty straightforwardly. To be precise, it is plausible that an author's scientific belief turns out false, which is the very reason why peer review is needed. Scientific consensus can be reached after evaluation not only by referees but also by peers in general (David, 1998). Thus, it is not that reporting untruth is a norm violation but that reporting what the author does not believe is. In an ideal process of peer review, authors should make every effort to convince referees of their version of truth. If authors spare this effort solely to avoid rejection or to expedite publication, the authors essentially send out a message that they do not really believe. That is, if an author revises his or her article by following referee's instruction, and if the revision ends in sending a message that is inconsistent with the author's scientific belief, we regard it as a breach of the honesty norm. Since referees and editors can also be responsible for dishonest conformity, we attempt to carefully control for relevant peer-review conditions, to highlight dishonesty on the authors' part, and to extrapolate our findings to scientific dishonesty in general.

4. Sample and Data

This study draws on empirical data in the field of life sciences in Japan. In Japan, research integrity has been on the important science policy agenda only for the past decade, although Japan is ranked third after the US and Germany in terms of the frequency of publication retractions (Fang et al., 2012). Matsuzawa (2013), analyzing misconduct reported by the Japanese media since the 1970s, shows that media exposure began increasing around 2000. Governmental surveys revealed that the Japanese science system had very poor preventive measures against scientific misconduct; for example, only 97 of 838 research associations (12%) had an ethics code before 2004, and only 11% of universities had any written misconduct policies before 2006 (Ishibashi and Ohtake, 2009). Responding to this situation, the government started serious debate, and the Ministry of Science finally published a guideline for research misconduct in 2006 (MEXT, 2006). Then, universities quickly set institutional guidelines for misconduct; as of 2008, 71% of universities had some guidelines (Ishibashi and Ohtake, 2009). Thus far, prevention and enforcement measures have been left up to individual institutions. Japan does not have an independent organization that specializes in research integrity, like ORI in the US, although the Science Council of Japan called for establishing one (SCJ, 2013).

In this study, we mainly use questionnaire data of Japanese life-science professors. We first conducted interviews of 21 principal investigators (PIs) of life science laboratories. Each interview took 1-2 hours. We investigated the interviewees' publication records in advance and asked a series of questions about their specific publications, such as their publication strategies and the impact of recent policies on the practice in publication. Based on the interview results, we prepared a survey instrument. The survey covers a series of questions regarding publication, from which this study uses the part related to peer review.

For the survey, we defined our sample as scientists (1) who published at least one paper in 2012-2013 in a journal registered in the Web of Science (WoS) in a field of life sciences as a corresponding author;[5] (2) who are affiliated with a life science department in a Japanese national university;[6] (3) who are PIs, and thus, have the authority in decision-making in publication, and (4) who are Japanese.[7] Since we did not have a comprehensive list of Japanese PIs, we first collected the information of WoS publication authors and then confirmed their organizational status using public information. We obtained 26,886 articles from the WoS and identified 13,877 unique corresponding authors satisfying conditions (1) and (2). We randomly sampled 1,700 authors with two strata of the Impact Factor (IF) of articles and the rank of the corresponding author's university. The journal IFs were divided into three strata, top10%, middle 70%, and bottom 20%. As for university ranks, the seven pre-imperial universities were defined as Tier 1;[8] the top half universities other than Tier 1, in terms of the publication count, was defined as Tier 2; and the rest were labeled Tier 3.[9] Then, we checked their organizational status as of 2013 and dropped those who did not satisfy the sampling criteria; i.e., those that were not a

[5] Fields are determined by WoS Subject Categories.
[6] The primary player in academic research in Japan is national universities. As of 2013, Japan has 86 national universities, of which 53 have departments related to life sciences (e.g., medicine, science, agriculture).
[7] We set this condition to spare the effort to prepare a questionnaire in foreign languages. Comparing foreign-borns and domestic scientists is of interest, but because the proportion of foreign-born scientists in Japan is extremely small (Franzoni et al., 2012), we supposed that meaningful comparison would be difficult.
[8] The top seven universities are designated as pre-imperial colleges and have been enjoying exceptionally prestigious status both in research and education (Kneller, 2007).
[9] We first grouped the authors into three scientific fields: medicine, agricultural sciences, and basic biology, based on WoS Subject Categories. Then, we employed stratified random sampling in each field. For the IFs, since we anticipated that authors of high-IF papers are relatively unwilling to respond for being busy, etc., we oversampled high-IF authors. In addition, we were interested in the behavior of low-IF authors, they were also oversampled. Precisely, sampling weights of 5.0, 1.0, and 2.5 are given to top, middle, and bottom IFs, respectively. For university ranks, sampling weights of 1.0, 1.5, and 2.0 are given to Tiers 1, 2, and 3, respectively. We oversampled lower-ranked universities because their population is smaller.

PI, were not affiliated with a life science department, were retired, or were non-Japanese. The final sample consisted of 777 PIs. We mailed the questionnaire to the PIs in November 2013, failing to reach six of them, and collected 358 responses (response rate = 46.4%) after two waves of requests until January, 2014.

The respondents were in the fields of basic biology (53%), medicine (26%), and agricultural sciences (21%); in the university strata of Tier 1 (33%), Tier 2 (32%), and Tier 3 (35%); and in the publication IF strata of the top (29%), middle (43%), and bottom (27%). Fifty-five percent were full professors, and 45% were associate professors. Six percent were female and 94% were male.

5. Description of Dishonest Conformity in Peer Review

In the survey, we needed to inquire into detailed settings of the peer review process, in which our respondent had to decide whether to reluctantly follow a referee's instructions or not. To this end, we focused on one incident of peer review for each respondent's peer review during the past one year as either a leading or corresponding author. Our respondents submitted, on average, 4.8 papers as a leading or corresponding author in one year, and were instructed to revise 3.2 papers upon peer review. Twenty-one percent of our respondents had no paper that needed revision and were dropped from the following analyses. The rest of the 284 respondents were asked if they had received a revision instruction that contradicted their scientific belief, and 136 answered that they had had such an experience at least once during the year. Thus, approximately half of our respondents who received revision instruction felt that the instruction was inconsistent with their belief. We then asked these respondents detailed questions about the peer review process for the specific instance.[10]

Reaction to revision instructions. We found that 63% of the respondents followed the inconsistent instructions, which we regard as dishonest conformity. On the other hand, 19% resubmitted the manuscript to the same journal without following the referees' instructions, 16% submitted the manuscript to a different journal without following the referees' instructions, and 2% gave up publishing the paper at all, all of which are regarded as not being dishonest conformity (Figure 1A).

Scientific value. Regardless of the respondents' reaction, we asked how they perceived the scientific value of the review comments by asking, 'How did you think it would affect the scientific value of your paper to follow the revision instructions completely?' with a five-point scale (1: negatively – 5: positively). The response was rather positive with 45% of the respondents positively appreciating the referees' instructions, but 15% perceived a negative impact, and 40% thought that revision would make no scientific difference (Figure 1B). If the perception of our respondents was right, this result implies that peer review improves the value of published findings for nearly half of cases, but that peer review wastes authors' time on other cases with either no meaningful impact or sometimes a negative impact. The reaction of the authors differed with this perceived value; when it was negative, the likelihood of dishonest conformity was lower but 47% still followed the referees' instructions.

Reason for inconsistent instruction. To examine possible reasons for the inconsistency, we asked, 'Why do you think the referees made such an instruction (one that is inconsistent with

[10] If they had multiple such instances, we asked them to choose one specific case that they found the most serious inconsistency.

your belief)?’ with three options. First, 49% of the respondents admitted that the authors themselves were responsible, for example, due to a lack of clarity. Second, 46% blamed the referees for a lack of understanding or relevant knowledge. Third, 15% answered that the reason was a competitive relationship between the referees and the authors. When the respondents admitted their own responsibility, they were more likely to follow the instructions. However, whether respondents blamed referees for the inconsistency did not affect their reaction. Though insignificant, a competitive relationship tended to decrease the likelihood of dishonest conformity.

Part of a paper to be revised. We inquired into what part of a paper the respondents were instructed to revise. The majority of our respondents received instructions to add data or experiments (68%) and to correct the interpretation of the results (60%), and only 12% of the respondents were instructed to change the hypotheses or research questions. Among the respondents who were instructed to add data or experiment, 36% considered that completely following the instruction was infeasible, and inevitably, tended not to conform to the instruction.

Impact of journals. We asked the name of the specific journal that gave the inconsistent instruction, but only 58% of our respondents answered this question, possibly because writing down a journal name is cumbersome or because they wanted to keep it confidential. With the available data, we analyze how journal prestige affects the decision of authors. To this end, we draw on Impact Factor (IF)[11] since IF has controversially prevailed as a journal ranking system in life sciences (Mcallister et al., 1980). A T-test shows that the mean IF for the group who followed a referee’s instruction is lower than that for the group who did not (4.0 vs. 7.4, $p < .1$). Thus, authors of low-impact journals seem more prone to dishonest conformity.

[11] We use JCR2012.

Result of peer review. Finally, we inquired into the result of peer review. In case our respondents did not follow the instructions and resubmitted the manuscript to the same journal, 70% of the respondents had their revision accepted, whereas 12% were instructed to revise once again and 18% had their papers rejected (Figure 1C). Thus, refusing to follow referees' instructions may be a risky option, but immediate rejection occurs only one out of five times. We also examined the ultimate result of peer review. While 93% of conformers succeeded in publishing their paper in the journals that gave inconsistent instructions, 48% of non-conformers eventually changed journals (Figure 1D).

<< Insert Figure 1 about here >>

6. Determinants of Conformity in Peer Review

To examine the determinants of dishonest conformity, we regress the dummy variable of dishonest conformity on the respondents' attributes and the peer-review conditions. Table 1 shows the descriptive statistics and correlation matrix of the variables. Table 2 presents the regression results of the following four models. Model 1 is the base model. Model 2 uses a subsample, dropping respondents who thought that completely following the referees' instruction was technically infeasible. Model 3 uses another subsample whose journal IF data is available. Model 4 adds a subjective value measurement (explained below) to a whole sample. Since the four models show qualitatively similar results, we only explain Model 1 unless we specify otherwise.

Additional measurements for the regressions are as follows. *MD* is a dummy variable assigned one if a respondent's first degree was from a medical school. *AP* is a dummy variable assigned one for associate professors and zero for full professors. *Age* is the age of respondents at the time of the survey. *Female* is a dummy variable assigned one for women and zero for men. As a performance measure, we summed up the citation counts of each respondent's papers published from 2009 to 2013 and took its logarithm (*ln(#Times cited)*). We inquired into the length of research experience in foreign countries with a six-point scale: 1) none, 2) half a year of less, 3) one year, 4) 2 years, 5) 3 years, and 6) 4 years or more (*foreign experience*). We asked the number of competitors in the respondents' research field with a five point scale: 1) none, 2) 1-2, 3) 3-5, 4) 6-10, and 5) 11 or more *(#competitors*). To measure subjective values about honesty in publication, we asked if the respondents agreed with the statement: 'Not reporting results inconsistent with author's claim is permissible to some extent, because it is difficult to explain complex natural phenomena perfectly consistently' with a five-point scale from 1) disagree to 5) agree (*selective publication*).

Peer-review condition. As for review conditions, technical feasibility is obviously a strong determinant ($b = -2.63$, $p < .001$). When authors positively appreciate the overall scientific value of a referee's comments, the likelihood of conformity is high ($b = .101$, $p < .05$). When respondents think that they are responsible for receiving inconsistent instructions, they are likely to conform (Model 4: $b = 1.37$, $p < .1$). When they feel that referees are responsible, instead, this decreases the likelihood of conformity though the effect is insignificant. Finally, when they perceive that competition with the referee led to inconsistent instructions, it also decreases the likelihood of conformity though insignificantly.

Scientific field. Comparing three scientific fields, basic biologists are the most susceptible to referees' instructions, and agricultural scientists are the least. The difference between the two fields is statistically significant ($b = -2.77$, $p < .05$). MD shows a significantly positive effect (b =

2.14, $p < .05$). Thus, while scientists in the field of medicine are not prone to dishonest conformity, scientists who received education in medical schools are. This suggests that the propensity to dishonest conformity may be determined in an early career stage.

Organizational rank. Associate professors are more prone to dishonest conformity than full professors ($b = 1.36$, $p < .1$). An interpretation for this result is that associate professors are under stronger pressure for publication, and thus, cannot afford to forgo the opportunity to publish by not following referees. Alternatively, associate professors may be less established and less confident in their scientific claims, and thus, more likely to yield to referees.

Performance. Performance measured by citation count shows a significantly negative effect ($b = -.62$, $p < .05$). Thus, high-performers are less likely to follow referees' comments than are low-performers. This can be understood in a similar way to the effect of organizational rank; high performers may not be desperate to publish many papers or may be confident enough not to adhere to the referees' instructions.

Competition. The number of competitors has a positive effect ($b = .67$, $p < .05$), straightforwardly suggesting that competitive pressure causes dishonesty. This is consistent with prior findings, suggesting that intense competition can lead to misconduct and questionable practices and jeopardize research integrity (Anderson et al., 2007c; De Vries et al., 2006).

Foreign experience. Foreign experience shows a strongly significant effect ($b = -.68$, $p < .01$), suggesting that returnees from foreign experience are less likely to engage in dishonest conformity. Examining in detail, we find that the propensity to conform is significantly lower for those who stayed abroad for one year or longer than for those who stayed for shorter terms. A few interpretations are plausible. First, returnees may be more confident in their scientific capability and handling of scientific debate with referees. Second, scientists expand their network through international experience, which might increase the likelihood of their dishonesty being uncovered.[12] Third, scientists who stayed abroad may have received better training for research integrity than is offered in Japan. As discussed above, the Japanese scientific community is lagged in terms of education for RCR.

Journal impact. When using journal IF as a continuous variable, we find no significant effect. However, categorizing low-IF (IF < 2), high-IF (IF > 8), and middle-IF (IF between 2 and 8),[13] we find that peer review in low-IF and high-IF journals is prone to dishonest conformity in comparison to that in middle-IF journals (Model 3: $b = 4.03$, $p < .05$; $b = 1.90$, $p > .1$, respectively). The positive coefficient of high-IF, though insignificant, is straightforward because the temptation for publication must be great. This is consistent with Van Noorden (2011), who suggests that retraction tends to happen in high-impact journals. Our interviewees also suggested that increasing emphasis on high-impact journals in evaluations had invited deceitful publications in top journals. Interestingly, our result suggests that conformity is significantly more likely to occur in low-IF journals. In fact, Van Noorden (2011) finds that the recent increase in retractions is attributed to low-IF journals. A few interpretations are plausible. For example, since low-IF journals do not attract as many readers, authors may think that dishonesty is unlikely to be uncovered or that possible damage to the scientific community by being dishonest should be limited. Alternatively, authors' motivation to publish in low-IF journals may be primarily to gain in publication metric but not to inform the scientific community, so they do not really care about what their paper claims. In any case, the result

[12] However, when we examined the effect of the size of network (e.g., the number of coauthors, etc.), we did not find a significant effect.
[13] We determined the threshold on the basis of our interviews.

seems to imply that the credibility of publications in low-IF journals can be compromised through peer review.

Subjective opinion. We additionally inquired into subjective opinions about scientific honesty in publication. To the question, 47% of the respondents disagreed while 14% agreed (Figure 2). We expected that this question correlates with the extent of honesty, and therefore, with the conformity in peer review. Confirming this idea, Model 4 shows a significantly positive coefficient of the opinion variable ($b = .64$, $p < .05$).

<< Insert Figure 2 about here >>

<< Insert Tables 1 & 2 about here >>

7. Discussion and Conclusion

Publication of scientific papers and intellectual honesty therein are critical for the progress of science (Merton, 1973; Zuckerman, 1977), but scientific dishonesty is commonly and increasingly observed in the forms of misconduct and other questionable practices (Fanelli, 2009; Fang et al., 2012; Martin, 2013; Martinson et al., 2005). Since this can seriously compromise the basis of the science system, understanding why scientists deviate from the honesty norm is crucial for any necessary action to be taken. However, empirical evidence has been rather limited, and the effect of extant countermeasures has been questioned (Anderson et al., 2007a). Among the various forms of dishonesty, this study focuses on one type of questionable practice, dishonest conformity in peer review; that is, authors compromise their own scientific belief amid the veto power of referees to have their papers published. Dishonest conformity has long been known (Frey, 2003). Although anecdotally it seems prevalent and the negative impact on science can be serious, empirical evidence has been lacking. This study investigates dishonest conformity based on survey data of Japanese life-science professors in an attempt to add to the literature on research integrity and on the peer review system.

Our results suggest that conflicts between authors and referees are quite common; approximately half of our respondents who received revision instructions during the recent one year felt that the instruction was inconsistent with their scientific belief. A majority of the instructions required adding data or experiments and correcting their interpretation of the results. While more than a few respondents positively perceived the overall value of review comments, the majority felt that the review instructions had either no scientific impact or a negative scientific impact. This result casts some doubt on the effectiveness of the current peer review system, and it calls for better incentive mechanisms for referees and careful monitoring by editors as prior literature implies.

Upon inconsistent instructions, 63% of respondents followed the instruction while 37% did not, suggesting that dishonest conformity is rather common. The payoff of dishonest conformity seems great, because 93% of the conformers successfully published their paper in the focal journal while only 52% of the non-conformers did; i.e., 48% of the non-conformers had to change journals. This difference may be too tempting to resist for authors who face inconsistent instructions. However, rebutting referees' argument is not entirely futile; when authors resubmit a revised paper to the same journal without following the inconsistent part of referees' instructions, it is only at the probability of 18% that the paper is immediately rejected. Though this result needs cautious interpretation due to its small sample size, it implies that the editors' monitoring may be functioning reasonably well.

Further, our results suggest some determinants of dishonesty by regression analyses of dishonest conformity. First, the propensity to conformity differs by field, consistent with Anderson et al. (1994). In particular, we find that basic biologists are more likely to conform than agricultural and medical scientists. This result is different from that of Fanelli (2009), who shows that misconduct is more frequently committed by medical and pharmacological researchers than others. Thus, depending on the types of dishonesty, the determinants can be different. The results also indicate that scientists who graduated from medical schools tend to conform. Thus, the trait of dishonesty may develop in early career stages, so training for research integrity should be given as early as possible. A comparison of organizational ranks suggests that associate professors are more likely to conform than full professors. Martinson et al. (2005), comparing mid- and early-career scientists, find that the former is more prone to questionable practices. Taken together, the propensity to dishonesty may be determined by the balance between pressure for publication, which is higher earlier in careers, and capabilities to circumvent honesty, which may be higher later in careers. We also find that scientists under competition are more likely to conform. This is consistent with De Vries et al. (2006), who also finds a positive correlation between questionable practices and competition, supporting that excessive pressure for publication can compromise research integrity. The results also suggest that scientists who have stayed abroad for one year or longer are less likely to conform. The background mechanism for this needs further investigation, but we suppose that a wider international network or advanced RCR education may discourage dishonesty. Finally, we find that conformity tends to occur in low-impact journals. Prior literature suggests that the emphasis on impact invites dishonesty in high-impact journals (Van Noorden, 2011), but our results suggest that different mechanisms, such as a weaker policing system, are at work for low-impact journals.

In sum, this study presents some empirical evidence for potential causes of scientific dishonesty; some of them are consistent with prior literature while others are inconsistent or newly suggested. Future research is thus needed to describe a more general picture of scientific dishonesty. Different types of dishonesty should be examined; as mentioned, background mechanisms can differ considerably. In addition, since this study draws on a specific sample of Japanese life scientists, studies in other countries and in other fields are needed. In fact, as for the country specificity, our results show a significant effect of foreign experience. Our results need to be interpreted with some reservations. For example, the measurement of dishonest conformity needs careful interpretation. It is self-reported, and desirability bias is possible. In addition, though we assume that dishonest conformity occurs only after referees provide comments, it is plausible that authors dishonestly prepare the first manuscript before submission, anticipating the likely instructions from referees (Frey, 2003). Therefore, a more comprehensive process of publication needs to be analyzed to fully understand where dishonesty is introduced into published findings. Furthermore, we focus on the dishonesty from the authors' side, but the dishonesty of referees also needs to be investigated.

References

Anderson, M. S., Louis, K. S., Earle, J. 1994. Disciplinary and departmental effects on observations of faculty and graduate student misconduct. ***Journal of Higher Education***, 65: 330-350.

Anderson, M. S., Horn, A. S., Risbey, K. R., Ronning, E. A., De Vries, R., Martinson, B. C. 2007a. What do mentoring and training in the responsible conduct of research have to do with scientists' misbehavior? Findings from a national survey of nih-funded scientists. ***Academic Medicine***, 82: 853-860.

Anderson, M. S., Martinson, B. C., De Vries, R. 2007b. Normative dissonance in science: Results from a national survey of us scientists. ***Journal of Empirical Research on Human Research Ethics***, 2: 3-14.

Anderson, M. S., Ronning, E. A., De Vries, R., Martinson, B. C. 2007c. The perverse effects of competition on scientists' work and relationships. ***Science and Engineering Ethics***, 13: 437-461.

Azoulay, P., Furman, J. L., Krieger, J. L., Murray, F. E. 2012. The effects of scientific paper retractions, ***NBER Working Paper***. Washington, D.C.: NEBR.

Bosch, X., Hernandez, C., Pericas, J. M., Doti, P., Marusic, A. 2012. Misconduct policies in high-impact biomedical journals. ***Plos One***, 7.

Broad, W. J. 1981. The publishing game - getting more for less. ***Science***, 211: 1137-1139.

Coupe, T. 2004. What do we know about ourselves? On the economics of economics. ***Kyklos***, 57: 197-215.

Csada, R. D., James, P. C., Espie, R. H. M. 1996. The "file drawer problem" of non-significant results: Does it apply to biological research? ***Oikos***, 76: 591-593.

David, P. A. 1998. Communication norms and the collective cognitive performance of "invisible colleges". In Navaretti, G. B. & Dasgupta, P. & Mäler, K.-G. & Siniscalco, D. (Eds.), ***Creation and transfer of knowledge: Institutions and incentives***. Heidelberg and New York: Springer.

De Vries, R., Anderson, M. S., Martinson, B. C. 2006. Normal misbehavior: Scientists talk about the ethics of research. ***Journal of Empirical Research on Human Research Ethics***, 1: 43-50.

Ellison, G. 2002a. Evolving standards for academic publishing: A q-r theory. ***Journal of Political Economy***, 110: 994-1034.

Ellison, G. 2002b. The slowdown of the economics publishing process. ***Journal of Political Economy***, 110: 947-993.

Fanelli, D. 2009. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. ***Plos One***, 4.

Fanelli, D. 2010. Do pressures to publish increase scientists' bias? An empirical support from us states data. ***Plos One***, 5.

Fang, F. C., Steen, R. G., Casadevall, A. 2012. Misconduct accounts for the majority of retracted scientific publications. ***Proceedings of The National Academy of Sciences of The United States of America***, 109: 17028-17033.

Franzoni, C., Scellato, G., Stephan, P. 2012. Foreign-born scientists: Mobility patterns for 16 countries. ***Nature Biotechnology***, 30: 1250-1253.

Frey, B. S. 2003. Publishing as prostitution? - choosing between one's own ideas and academic success. ***Public Choice***, 116: 205-223.

Grieneisen, M. L., Zhang, M. 2012. A comprehensive survey of retracted articles from the scholarly literature. ***Plos One***, 7: 1-15.

Hagstrom, W. O. 1974. Competition in science. ***American Sociological Review***, 39: 1-18.

Hess, D. J. 1997. ***Science studies: An advanced introduction***. New York: New York University Press.

Ietto-Gillies, G. 2012. The evaluation of research papers in the xxi century. The open peer discussion system of the world economics association. ***Frontiers in Computational Neuroscience***, 6.

Ishibashi, E., Ohtake, Y. 2009. Report of the survey of scientific, technological and academic activities in the universities, ***NISTEP Report***, Vol. 167.

John, L. K., Loewenstein, G., Prelec, D. 2012. Measuring the prevalence of questionable research practices with incentives for truth telling. ***Psychological Science***, 23: 524-532.

Macrina, F. L. 2007. Scientific societies and promotion of the responsible conduct of research: Codes, policies, and education. ***Academic Medicine***, 82: 865-869.

Martin, B. R. 2013. Whither research integrity? Plagiarism, self-plagiarism and coercive citation in an age of research assessment. ***Research Policy***, 42: 1005-1014.

Martinson, B. C., Anderson, M. S., De Vries, R. 2005. Scientists behaving badly. ***Nature***, 435: 737-738.

Martinson, B. C., Anderson, M. S., Crain, A. L., De Vries, R. 2006. Scientists' perceptions of organizational justice and self-reported misbehaviors. ***Journal of Empirical Research on Human Research Ethics***, 1: 51-66.

Marusic, A., Bosnjak, L., Jeroncic, A. 2011. A systematic review of research on the meaning, ethics and practices of authorship across scholarly disciplines. ***Plos One***, 6.

Matsuzawa, T. 2013. Research misconduct in japan: Macro-analysis based on open information. ***Journal of Information Processing and Management***, 56: 156-165.

Mcallister, P. R., Anderson, R. C., Narin, F. 1980. Comparison of peer and citation assessment of the influence of scientific journals. ***Journal of the American Society For Information Science***, 31: 147-152.

Merton, R. K. 1973. ***Sociology of science***. Chicago: University of Chicago Press.

MEXT. 2006. Guideline for measures against research misconduct. Tokyo: MEXT.

NAS. 1992. Responsible science, volume i: Ensuring the integrity of the research process. Washington, D.C.: NATIONAL ACADEMY PRESS.

Pitsoulis, A., Schnellenbach, J. 2012. On property rights and incentives in academic publishing. ***Research Policy***, 41: 1440-1447.

Prufer, J., Zetland, D. 2010. An auction market for journal articles. ***Public Choice***, 145: 379-403.

Resnik, D. B., Master, Z. 2013. Policies and initiatives aimed at addressing research misconduct in high-income countries. ***Plos Medicine***, 10.

SCJ. 2013. Preventive and ex-post measures against research misconduct. Tokyo: SCJ.

Tsang, E. W. K., Frey, B. S. 2007. The as-is journal review process: Let authors own their ideas. ***Academy of Management Learning & Education***, 6: 128-136.

Van Noorden, R. 2011. The trouble with retractions. ***Nature***, 478: 26-28.

Wilhite, A. W., Fong, E. A. 2012. Coercive citation in academic publishing. ***Science***, 335: 542-543.

Zuckerman, H. 1977. Deviant behavior and social control in science. In Sagarin, E. (Ed.), ***Deviance and social change***. Beverly Hills: Sage Publications.

Table 1 Descriptive statistics and Correlation Matrix [a]

Variable	Mean	S.D.	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21
1. Dishonest conformity	.62	.49	.00	1.00																					
2. Technical infeasibility	.27	.44	.00	1.00	***-.45***																				
3. Scientific value	3.28	.89	1.00	5.00	***.23***	.02																			
4. Referee responsible	.46	.50	.00	1.00	***-.20***	.01	***-.37***																		
5. Author responsible	.48	.50	.00	1.00	***.39***	-.17	***.35***	***-.32***																	
6. Low IF	.24	.43	.00	1.00	.16	.03	.01	-.01	.19																
7. Middle IF	.60	.49	.00	1.00	-.05	-.22	-.09	.09	-.12	***-.68***															
8. High IF	.16	.37	.00	1.00	-.11	***.26***	.11	-.11	-.06	***-.24***	***-.54***														
9. Tier 1 Univ	.30	.46	.00	1.00	-.08	.06	-.11	.09	-.16	-.22	.09	.14													
10. Tier 2 Univ	.33	.47	.00	1.00	.12	.01	.17	-.08	.09	.04	-.10	.08	***-.46***												
11. Tier 3 Univ	.37	.48	.00	1.00	-.04	-.07	-.07	-.01	.06	.17	.01	-.21	***-.50***	***-.54***											
12. Biology	.16	.37	.00	1.00	-.17	.15	.07	.15	-.05	.20	-.10	-.10	-.13	-.11	***.24***										
13. Agriculture	.51	.50	.00	1.00	.12	-.12	-.08	-.13	.02	-.14	.24	-.15	***.24***	-.12	-.12	***-.44***									
14. Medicine	.33	.47	.00	1.00	.01	.01	.03	.02	.02	-.01	-.17	.23	-.15	***.21***	-.06	***-.31***	***-.72***								
15. MD	.28	.45	.00	1.00	.14	-.03	-.05	-.04	.00	-.24	.05	.21	-.16	.08	.07	***-.27***	***-.44***	***.67***							
16. AP	.41	.49	.00	1.00	***.24***	-.12	-.09	.02	***.21***	.10	-.05	-.04	-.06	***.17***	-.10	-.13	-.06	.17	.06						
17. Age	52.19	7.25	38.00	66.00	-.02	.02	.11	-.08	-.01	-.05	-.06	.14	.14	***-.24***	.10	.16	-.08	-.03	.02	***-.51***					
18. Female	.07	.25	.00	1.00	-.06	.07	-.12	-.04	.01	-.02	.00	.03	-.10	.17	-.06	-.12	.06	.02	-.02	***.19***	-.12				
19. ln(#Times cited)	4.13	1.47	.00	7.27	-.06	-.07	.08	.00	-.11	***-.34***	.14	.21	***.31***	-.01	***-.28***	***-.41***	.05	***.27***	***.31***	-.17	-.04	-.06			
20. #Competitors	2.90	1.21	1.00	5.00	.12	-.03	-.05	.01	.02	-.22	.11	.11	-.10	***.22***	-.12	.05	-.04	.00	***.19***	.00	-.03	.16	.16		
21. Foreign experience	2.25	1.52	.00	5.00	-.14	.10	.16	-.04	-.04	***-.26***	.06	.21	-.08	.11	-.02	-.18	-.02	.16	***.26***	-.10	.07	-.02	.16	***.20***	
22. Selective publication	2.53	1.07	1.00	5.00	.14	.09	.01	.01	-.01	.03	-.11	.10	-.11	.01	.09	.06	***-.21***	***.18***	***.25***	.03	.07	.02	.01	.03	.05

[a] Bold italic: $p < .05$. N = 120 (except for variables 6-8 with N =68).

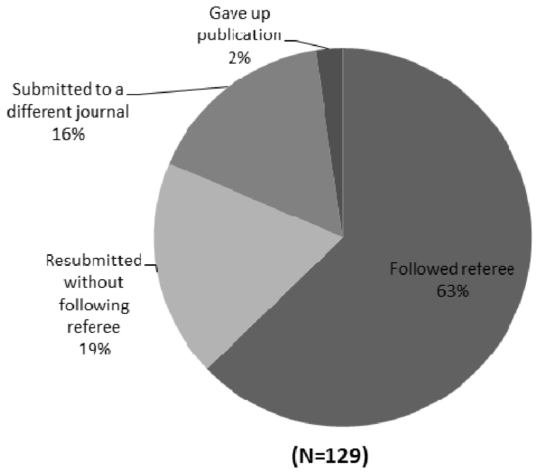
Table 2 Prediction of Dishonest Conformity [a]

	Model 1 Whole sample		Model 2 Technically feasible		Model 3 IF available		Model 4 Whole sample	
Technical infeasibility	-2.63 ***	(.69)			-5.55 **	(2.14)	-3.04 ***	(.78)
Scientific value	1.01 *	(.40)	1.57 **	(.58)	1.01	(.80)	.99 *	(.41)
Author responsible	1.13	(.70)	.99	(.90)	1.61	(1.41)	1.37 †	(.75)
Referee responsible	-.16	(.69)	-.38	(1.00)	-2.26	(1.50)	-.18	(.71)
Referee competes with authors	-.52	(.90)	-2.17 †	(1.26)	-.93	(1.38)	-.37	(.94)
Low IF					4.03 *	(1.93)		
Middle IF (base group)								
High IF					1.90	(1.69)		
Age	.06	(.05)	.14 †	(.08)	.17	(.12)	.05	(.06)
Female	-1.84	(1.20)	-2.37	(2.13)	-3.91	(2.54)	-1.70	(1.27)
Tier 1 Univ (base group)								
Tier 2 Univ	.40	(.91)	.79	(1.28)	3.12	(1.93)	.16	(.96)
Tier 3 Univ	-.27	(.84)	1.26	(1.37)	.81	(1.45)	-.73	(.90)
Biology (base group)								
Agriculture	-2.77 *	(1.11)	-4.20 *	(1.95)	-5.87 *	(2.87)	-2.82 *	(1.11)
Medicine	-1.51	(.92)	-1.48	(1.14)	-6.70 **	(2.55)	-1.52	(.97)
MD	2.14 *	(1.00)	.96	(1.19)	4.96 *	(1.97)	1.95 †	(1.05)
AP	1.36 †	(.77)	3.53 **	(1.31)	.85	(1.89)	1.15	(.80)
ln(#Times cited)	-.62 *	(.27)	-.45	(.45)	-1.01	(.65)	-.75 *	(.29)
#Competitors	.67 *	(.30)	.96 *	(.43)	1.35 *	(.69)	.72 *	(.30)
Foreign experience	-.68 **	(.25)	-.98 *	(.42)	-1.19 *	(.59)	-.67 **	(.24)
Selective publication							.64 *	(.30)
χ^2	73.26 ***		47.53 ***		56.35 ***		78.53 ***	
Pseudo R^2	.459		.480		.598		.492	
N	120		88		68		120	

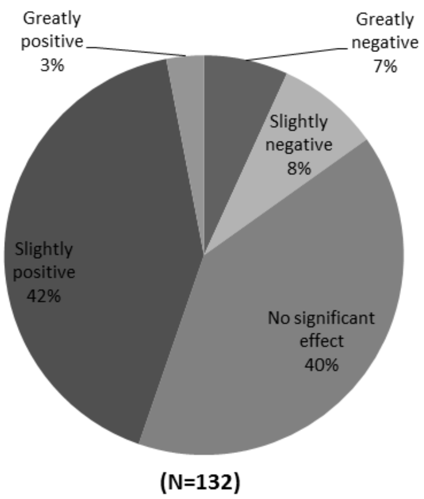
[a] Logit regressions. Unstandardized coefficients and standard errors (parentheses). Two-tailed test. † p<0.10; * p<0.05; ** p<0.01;*** p<0.001.

Figure 1 Description of Dishonest Conformity [a]

(A) Reaction to inconsistent instruction

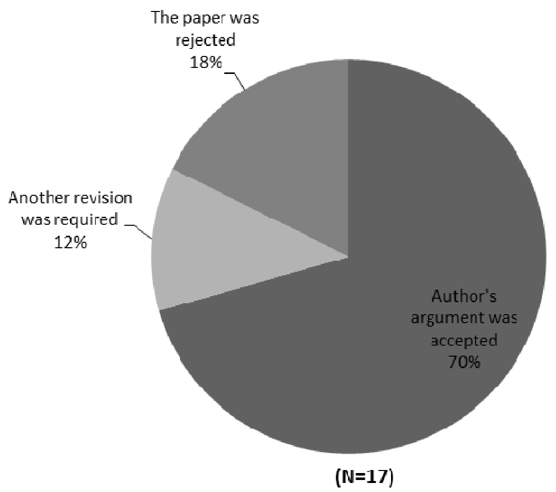


(B) Scientific value-added of overall review comments



[a] From (C) and (D) are dropped non-respondents and respondents whose review result had not been decided by the time of the survey.

(C) Result of not following referees



(D) Final result of peer review

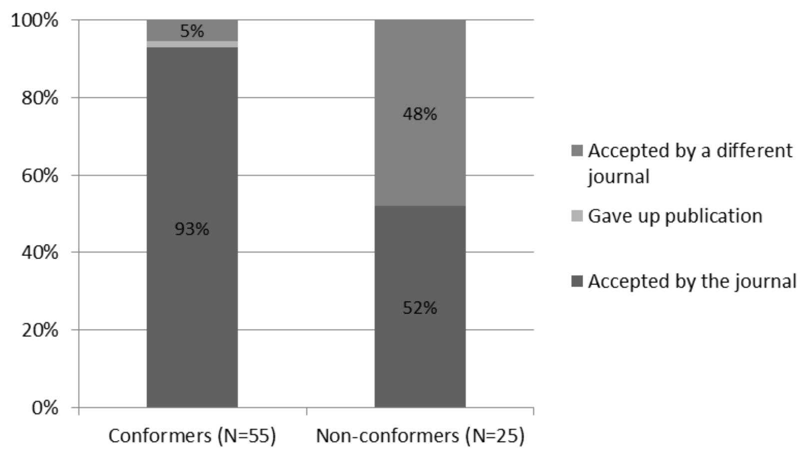


Figure 2 Opinions about intellectual honesty