Journal of Empirical Research on
Human Research Ethics
2014, Vol. 9(5) 64–71

DOI: 10.1177/1556264614552421
jre.sagepub.com

Publication Pressure and Scientific Misconduct in Medical Scientists

Joeri K. Tijdink[1,2], Reinout Verbeke[3], and Yvo M. Smulders[1]

Abstract

There is increasing evidence that scientific misconduct is more common than previously thought. Strong emphasis on scientific productivity may increase the sense of publication pressure. We administered a nationwide survey to Flemish biomedical scientists on whether they had engaged in scientific misconduct and whether they had experienced publication pressure. A total of 315 scientists participated in the survey; 15% of the respondents admitted they had fabricated, falsified, plagiarized, or manipulated data in the past 3 years. Fraud was more common among younger scientists working in a university hospital. Furthermore, 72% rated publication pressure as "too high." Publication pressure was strongly and significantly associated with a composite scientific misconduct severity score.

Keywords

scientific misconduct, publication pressure, fraud, questionable research practice, ethics in publishing

Recent years have witnessed increased attention for scientific misconduct (Couzin-Frankel, 2013; Fanelli, 2013; Wicherts, 2011). In 2011, a systematic review concluded that almost 2% of scientists confessed having fabricated or falsified data at least once, and up to 33% admitted to other "questionable research practices" (QRP; Fanelli, 2009). For scientific misconduct observed in colleagues, corresponding rates were no less than 14% and 72%, respectively (Lafollette, 2000). Although some scientists see fraudulent colleagues as "just a few bad apples," there is increasing evidence that scientific misconduct occurs on a scale that compromises the credibility of science.

There are different definitions and classifications for scientific misconduct. Fabrication, falsification, and plagiarism are usually qualified as fraud. Other actions that violate traditional values of the research process and may be detrimental to its credibility may be referred to as QRP (Steneck, 2006; Martinson, Anderson, & De Vries, 2005), typical examples of which include salami slicing, guest authorships, or intuitively deleting data.

Particularly in medicine, concerns have been expressed that scientists are continuously producing "publishable" results at the expense of quality, scientific rigor, and personal integrity (Fanelli, 2010; Tijdink, Vergouwen, & Smulders, 2013; Anderson, Ronning, De Vries, & Martinson, 2007). Excessive emphasis on scientific output may also jeopardize academic activities that compete with science for time and attention, such as clinical and educational duties (Deangelis, 2004).

Recently, we found that the degree to which medical professors in the Netherlands experience publication pressure correlates strongly with symptoms of job-related stress and burnout (Tijdink et al., 2013). Intuitively, publication pressure may also be a risk factor for scientific misconduct, but to our knowledge, this has not been studied. The main aim of this study is to address the potential relationship between publication pressure, and self-reported fraud and QRP.

Method

Procedure and Participants

All researchers working in medicine at one of the five academic medical centers in Flanders, Belgium, were sent an invitational e-mail in October 2012. The letter was distributed among the scientists by the communication office of the participating universities for privacy reasons. The invitational e-mail explained the objectives of the study (research of scientific culture) and provided them with a link to an anonymous online questionnaire on a protected website. All addressees were sent a reminder after 2 weeks.

[1]VU University Medical Center, Amsterdam, The Netherlands
[2]Tergooi Hospitals, Hilversum, The Netherlands
[3]*EOS* Magazine, Antwerp, Belgium

Corresponding Author:
Joeri K. Tijdink, VU University Medical Centre, Department of Internal Medicine, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands.
Email: j.tijdink@vumc.nl

Table 1. Publication Pressure Questionnaire (PPQ).

Questions (domain)	Likert-type scale score (*SD*)	% agreement (4-5 or 1-2[a] on Likert-type scale)
1. Without publication pressure, my scientific output would be of higher quality	2.9 (1.2)	35
2. My scientific publications contribute to better (future) medical care[a]	3.7 (0.9)	67
3. I experience my colleagues' assessment of me on the basis of my publications as stressful	3.4 (1.2)	52
4. I experience the publication criteria formulated by my university for my appointment or re-appointment as professor as a stimulus[a]	2.8 (1.0)	26
5. Publication pressure puts pressure on relationships with fellow researchers	3.5 (1.1)	59
6. I suspect that publication pressure leads some colleagues (whether intentionally or not) to color data	3.7 (1.1)	64
7. The validity of medical world literature is increased by the publication pressure in scientific centers[a]	2.2 (0.9)	11
8. Publication pressure leads to serious worldwide doubts about the validity of research results	3.6 (1.0)	61
9. In my opinion, the pressure to publish scientific articles has become too high	3.9 (1.0)	72
10. The competitive scientific climate stimulates me to publish more[a]	3.2 (1.0)	48
11. My colleagues judge me mainly on the basis of my publications	3.2 (1.2)	43
12. Fellow scientists maintain their clinical and teaching skills well, despite publication pressure[a]	3.1 (1.0)	39
13. I cannot confide innovative research proposals to my colleagues	2.5 (1.1)	20
14. Publication pressure harms science	3.4 (1.2)	52

[a]Inversed questions; higher scores for disagreement.

Variables

The questionnaire contained, apart from demographic data, two parts: a validated Publication Pressure Questionnaire (PPQ; Tijdink, Smulders, Vergouwen, de Vet, & Knol, 2014) and a 12-item questionnaire assessing scientific misconduct.

The validated PPQ contained 14 statements (Table 1), the responses to which were scored on a 5-point Likert-type scale. The Likert-type scale scores were assigned one to five points such that higher scores reflected higher pressure. We also designed five inversed questions within the questionnaire in which higher scores represented lower experienced publication pressure. We labeled different statements as positive and negative to avoid "yeah-saying" (Streiner & Norman, 2008). Negative statements were scored inversely (see Table 1).

The questionnaire assessing scientific misconduct consisted of 12 different types of scientific misconduct (see Table 2). Survey respondents were asked to report in each case whether or not (yes or no) they had themselves engaged in specified types of misconduct during the past 3 years, as well as to report whether they had personally witnessed colleagues showing these types of misconduct. The content of the questionnaire and severity scoring of misconduct was based on previous studies by Fanelli (2009) and Martinson (2005). For better understanding and differentiating fraud (fabrication, falsification, and plagiarism) from other types of scientific misconduct, the 12 items were divided into three categories: fraud (Items 1, 2, 7, and 11), severe scientific misconduct (Items 3, 4, 5, 8, 9, and 12), and "moderate" scientific misconduct (Items 6 and 10).

To construct a self-reported scientific misconduct severity index, the positive answers of the Fraud questions (Items 1, 2, 7, and 11) were assigned three points, positive answers of the severe scientific misconduct questions were assigned two points, and positive answers of the moderate scientific misconduct questions were assigned one point. Scores were added up to calculate a total scientific misconduct severity score (maximum range = 0-28).

Demographics

Respondents provided demographic information on gender, age, type of specialty; years working as a scientist; appointment status; main professional activity (research, education, patient care, or management); and Hirsch index.

Statistical Analysis

ANOVA was used to compare groups. Pearson's correlation coefficients were calculated to examine relationships between continuous variables. The *t* tests, ANOVAs, and univariate and multivariate linear regression analyses were

Table 2. From Fraud to Questionable Research Practices: 12 Types of Scientific Misconduct.

Scientific Misconduct Questionnaire: 3 years retrospectively (*n* = 315)	Yes % (*n*)
Fraud (fabrication, falsification, plagiarism)	
1a Have you fabricated data 1 or more times?	1% (4)
1b Have you observed that a colleague fabricated data?	24% (76)
2a Have you selectively deleted data to confirm a hypothesis ("massaging data," "cooking data")	7% (23)
2b Have you observed that a colleague selectively deleted data to confirm a hypothesis ("massaging data," "cooking data")	44% (138)
7a Have you ever modified the results of a study under pressure from an organization that funded the research?	4% (13)
7b Have you observed that a colleague modified the results of a study under pressure from an organization that funded the research?	18% (56)
11a Have you used the same data or results for two or more publications in different peer-reviewed journals (self-plagiarism)?	5% (15)
11b Have you observed that a colleague used the same data or results for two or more publications in different peer-reviewed journals (self-plagiarism)?	40% (126)
Severe misconduct	
3a Have you ever deleted observations or data from analyses because your intuition told you that they were incorrect?	26% (83)
3b Have you observed that a colleague deleted observations or data from analyses because his or her intuition told him or her they were incorrect?	43% (135)
4a Have you ever concealed data that contradict your previous research?	3% (10)
4b Have you observed that a colleague concealed data that contradict his or her previous research?	22% (70)
5a Have you ever used the ideas of others without their permission or without proper citation?	1% (2)
5b Have you observed that a colleague used the ideas of others without their permission or without proper citation?	36% (112)
8a Have you ever not published results under pressure from an organization that funded the research?	2% (6)
8b Have you observed that a colleague did not publish results under pressure from an organization that funded the research?	12% (39)
9a Have you ever deliberately not mentioned an organization that funded your research (that had a specific interest in the research) in the publication of your study?	0%
9b Have you observed that a colleague deliberately not mentioned an organization that funded your research (that had a specific interest in the research) in the publication of your study?	7% (21)
12a Have you ever been pressured into questionable research practices?	11% (33)
12b Have you observed that a colleague been pressured into questionable research practices?	19% (59)
Moderate misconduct	
6a Have you ever turned a blind eye to other people's use of flawed data or questionable interpretation of data?	20% (62)
6b Have you observed that a colleague turned a blind eye to other people's use of flawed data or questionable interpretation of data?	37% (116)
10a Have you ever (whether under pressure or not) added one or more authors to a study who did not contribute to the research?	69% (218)
10b Have you observed that a colleague (whether under pressure or not) added one or more authors to a study who did not contribute to the research?	85% (268)

used to identify confounders and potential mediators of determinants of QRP and PPQ scores. In the multivariate analysis, we entered the variables with p values of < .10 in the univariate analysis. The Statistical Package for the Social Sciences (SPSS) statistics (Chicago, USA 2011, Version 20) was used for the statistical analyses.

Results

According to the data provided by the administrative offices of the five participating Flemish medical universities, we reached approximately 2,548 scientists. Of the 2,548 scientists, there were 484 who responded (19%), of whom 315 (12%) completed the two questionnaires. The demographic data of the complete responders are summarized in Table 3.

Scientific Misconduct

Table 4 shows the percentages of respondents who reported that they had engaged in each type of behavior (from moderate scientific misconduct to fraud). Fifteen percent of respondents admitted that they had fabricated,

Table 3. Demographics.

Demographics	n = 315	%
Gender		
Male	170	54
Female	145	46
Age		
26-35	129	41
36-45	78	25
46-55	77	24
56-65	27	9
65 and older	4	1
Connected to university hospital		
Yes	175	56
No	140	44
Position		
PhD student	75	24
Postdoc	141	45
Professor	99	31
Years active in science		
0-5	82	26
6-10	73	23
11-15	64	20
15 or more	96	31
No. 1 work priority		
Research	184	58
Education	36	11
Patient care	75	24
Management	20	6
Appointment		
Temporary	165	52
Permanent	150	48

falsified, plagiarized, or manipulated data in the past 3 years (see Figure 1). More than one out of four ever deleted data or results to confirm a hypothesis (data cooking/massaging), and almost 70% of respondents assigned authorships to people did not contribute to the study (see Table 4 and Figure 1).

PPQ

Table 3 lists the questions and responses of the PPQ. The responses to some of the key questions indicate that 72% (Item 13) rated publication pressure as "too high." In addition, more than half of the respondents (64% and 61%; Items 7 and 8) believed that the pressure to publish has detrimental effects on the validity and credibility of medical science. Finally, 52% (Item 14) stated that publication pressure has a "harmful" effect on medical science.

We use the sum score of the 14 items (M = 31.26, SD = 8.56; inversed items taken into account) of the condensed, validated questionnaire for further analysis of correlations and determinants.

Correlation of Self-Reported Scientific Misconduct With Demographic Characteristics

Table 4 provides regression analysis results of demographic and job-specific variables as independent and total scientific misconduct severity score as the dependent variable. In univariate analyses, younger age and being an early career scientist were significantly associated with higher self-reported scientific misconduct scores. There were trends for lower misconduct scores among those with a fixed job position and those who were not connected to a university clinic. Entering these four variables in a multivariate analysis suggested that being connected to a university clinical hospital was the only independent correlate of self-reported scientific misconduct. The Hirsch index was not significantly correlated with fraud or QRP.

Scientific Misconduct and Publication Pressure

In both univariate analyses, as well as after adjustment for demographic variables, the PPQ score was strongly and

Table 4. Univariate and Multivariate Analysis Comparing Independent Variables With Total Scientific Misconduct Score.

	β	CI 95% lower bound	CI 95% upper bound	*p* value
Univariate				
Gender (female)	−.35	−0.84	0.14	.17
Connected to university hospital	.42	−0.69	0.92	.09
Fixed position (yes)	−.48	−0.97	0.11	.06
Age (four groups, increase of 5 years)	−.38	−0.61	−0.14	.002
Position (three groups)	−.23	−0.57	0.10	.17
Years active in science (increase of 5 years)	−.32	−0.52	−0.11	.003
PPQ	.07	0.04	0.10	.000
H index (*n* = 121)	−.01	−0.05	0.02	.45
Multivariate				
Model 1				
Connected to university hospital	**.51**	**0.01**	**1.01**	**.05**
Fixed position	.01	−0.66	0.68	.97
Age	−.32	−0.71	0.08	.12
Years active in science	−.11	−0.47	0.25	.55
Model 2				
Connected to university hospital	**.59**	**0.11**	**1.07**	**.02**
Fixed position	.07	−0.58	0.72	.83
Age	−.25	−0.63	0.14	.20
Years active in science	−.15	−0.49	0.20	.41
PPQ	**.07**	**0.04**	**0.10**	**.000**

Note. Model 1: Demographic and job-specific characteristics; Model 2: Add adjustment for PPQ. PPQ = Publication Pressure Questionnaire. Bolded variables have a significance level of $p < .05$.

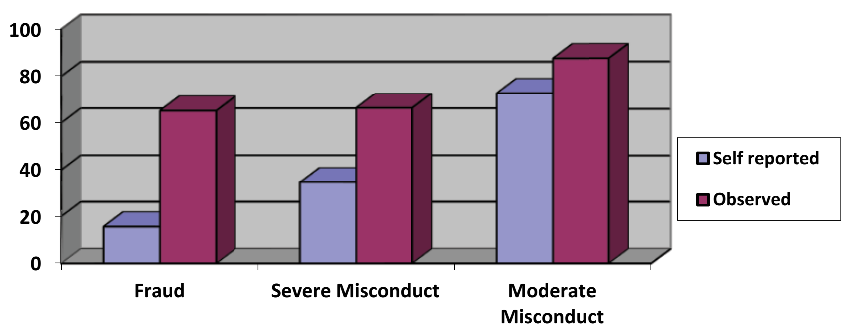

Figure 1. Percentages of scientific misconduct.

significantly associated with the scientific misconduct severity score. Figure 2 illustrates this correlation.

Discussion

This is the first study to address the possible association between publication pressure and (self-reported) scientific misconduct. The results support the notion that excessive pressure to publish scientific articles contributes to scientific misconduct, at least among European medical scientists.

Our findings are generally in line with one study published on this topic. This study has related publication pressure to publication bias (Fanelli, 2010). Our results add to previous literature in that our survey includes a validated measurement instrument of publication pressure together with a questionnaire exploring severity and different types of scientific misconduct. Both instruments were used in a relatively small, coherent community of five academic hospitals in a distinct European region (Flanders, Belgium).

The level of publication pressure reported by Flemish scientists can only be compared with the first study using this same questionnaire, which was performed in 2011 among Dutch medical professors (Tijdink et al., 2013). Insofar as this comparison is valid, publication pressure

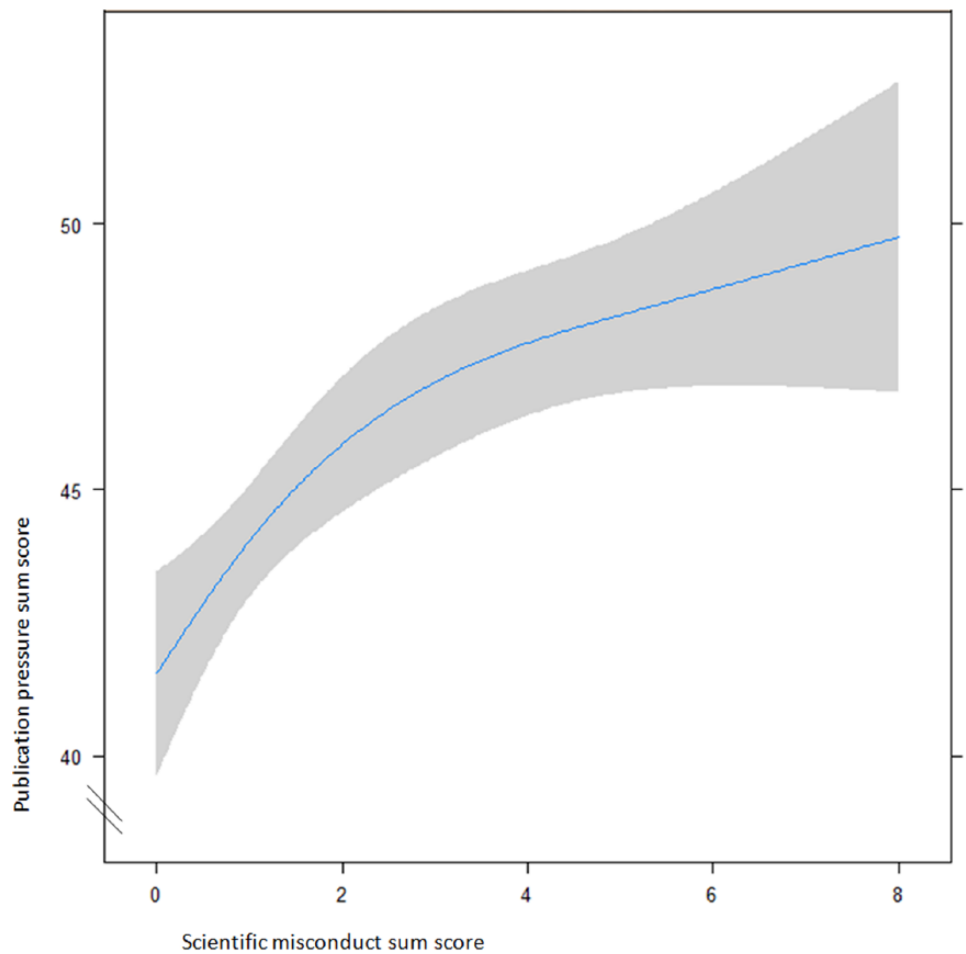

Figure 2. Cubic spline, demonstrating the association between the sum score on the validated 14-item Publication Pressure Questionnaire, and the scientific misconduct sum score (Wells, 1998).

appears higher in Flemish scientists (PPQ score 31.3 vs. 25.5 points).

Publication pressure may be considered a form of psychological stress. Stress is often seen as an inducer of risky behavior (Pabst, Brand, & Wolf, 2013; Vinkers et al., 2013). If we translate this to medical researchers, pressure and stress may contribute to scientific misconduct. It is conceivable that pressure generates stress which can influence the amount of errors made in scientific research. Such a relationship has been found among clinical doctors, where burnout is associated with higher prevalence of medical errors made (West, Tan, Habermann, Sloan, & Shanafelt, 2009).

An important cause of publication pressure could be the increasing importance of bibliographic parameters in the assignment of research funding, grants scholarships, and academic positions (Bedeian, Van Fleet, & Hyman, 2009; Bird, 2006; Gannon, 2000; Hessels, 2010). Although our study was not designed to identify causes of publication pressure, the responses to Statements 3, 4, and 10 suggest an important role of quantitative parameters, such as cumulative scientific output. In this respect, there is an interesting study of 331 biomedical postdoctoral fellows of whom only 3.4% admitted to having modified data in the past, but of whom 17% expressed they would be "willing to select or omit data to improve their results" (Eastwood, Derish, Leash, & Ordway, 1996). In another study of 549 biomedical trainees on the ethics of scientific investigation, 4.9% of biomedical early career scientists said they had modified research results in the past, but 81% were willing to omit or fabricate data to win a grant or publish an article (Kalichman & Friedman, 1992).

The interpretation of our results is limited by the possibility of substantial response bias. Hence, our data are more suited for identification of potential determinants of self-reported misconduct than they are for absolute prevalence of misconduct. Nonetheless, a few comparisons with existing literature are worth consideration. The self-reported rates of fraud and QRP are high compared with other surveys (Fanelli, 2009; Fanelli, 2009; Martinson et al., 2005). The survey performed by Martinson et al. (2005) on 3,247 scientists reported high percentages of different aspects of scientific misconduct, that is, inappropriately assignment of authorship credit (10%), withholding details of methodology or results in articles or proposals (14%), and deletion of observations or data points from analyses based on a "gut feeling" of inadequacy (15%). Fanelli has provided a pooled weighted average of 21 studies, reporting that 2% of scientists admitting to have fabricated, falsified, or modified data or results at least once, and up to 33% admitting other QRP. Corresponding rates for the behavior of colleagues were 14% for falsification and up to 72% for other QRP (Fanelli, 2009). Our study suggests a data fabrication, falsification, and plagiarism rate of 8%, which is in line with existing literature.

Limitations

First, with a response rate of 19% and a completion rate of 12%, response bias is likely. This will mainly affect reported publication pressure and prevalence rates of scientific misconduct but is less likely to cause Type I errors in associations between the two. The accuracy of responses to sensitive questions such as questions on manipulating data or fraudulent behavior is often independent of response rates and depends strongly on respondents' perception of anonymity and confidentiality (Bates & Cox, 2008; Tourangeau & Smith, 1996). Another potential source of bias, in either direction, could be the use of an online questionnaire. The validity of online questionnaires has been extensively studied, and there is no evidence that web-based questionnaires are less valid than "live" questionnaires (Joinson, 1999).

Phrasing of questions is also relevant. Scientists are, for example, less likely to reply positively to questions using the words "fabrication and falsification" rather than "alteration or modification." In our questionnaire, we not only used fabrication but also more positively phrased items such as "deletion of data because your intuition told you they were wrong."

We also considered a possible "Muhamed Ali effect," in which people perceive themselves as more honest than their peers. Researchers might be overindulgent with their own behavior and overzealous in judging their colleagues (Fanelli, 2009). Questions regarding colleagues' behavior might tend to inflate estimates of misconduct because the same incident might be reported by more than one respondent.

Research Agenda

This is the first article to address the potential effect of publication pressure on scientific misconduct among scientists in Flanders. Replication of this finding, preferably in collaborative efforts, is imperative, and can significantly impact scientific practice and policy making in other countries.

Educational Implications

The article can be used in educational settings. It is thought-provoking in terms of how to view and interpret science, scientists, and publication culture.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: We would like to thank the Pascal Decroos Foundation and ZonMw (The Netherlands Organisation for Health Research and Development) for supporting this research.

References

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

Bates, S. C., & Cox, J. M. (2008). The impact of computer versus paper-pencil survey, and individual versus group administration, on self-reports of sensitive behaviors. *Computers in Human Behavior*, *24*, 903-916.

Bedeian, A. G., Van Fleet, D. D., & Hyman, H. H. (2009). Scientific achievement and editorial board membership. *Organizational Research Methods*, *12*, 211-238.

Bird, S. J. (2006). Research ethics, research integrity and the responsible conduct of research. *Science and Engineering Ethics*, *12*, 411-412.

Couzin-Frankel, J. (2013). Shaking up science. *Science*, *339*, 386-389.

Deangelis, C. D. (2004). Professors not professing. *Journal of the American Medical Association*, *292*, 1060-1061.

Eastwood, S., Derish, P., Leash, E., & Ordway, S. (1996). Ethical issues in biomedical research: Perceptions and practices of postdoctoral research fellows responding to a survey. *Science and Engineering Ethics*, *2*, 89-114.

Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS ONE*, *4*, Article e5738.

Fanelli, D. (2010). Do pressures to publish increase scientists' bias? An empirical support from US States Data. *PLoS ONE*, *5*, Article e10271.

Fanelli, D. (2013). Redefine misconduct as distorted reporting. *Nature*, *494*, 149.

Gannon, F. (2000). The impact of the impact factor. *EMBO Rep*, *1*, 293.

Hessels, L. K. (2010). *Science and the struggle for relevance*. Utrecht University, Utrecht.

Joinson, A. (1999). Social desirability, anonymity, and Internet-based questionnaires. *Behavioral Research, Methods, Instruments & Computers*, *31*, 433-438.

Kalichman, M. W., & Friedman, P. J. (1992). A pilot study of biomedical trainees' perceptions concerning research ethics. *Academic Medicine*, *67*, 769-775.

LaFollette, M. C. (2000). The evolution of the "scientific misconduct" issue: An historical overview. *Proceedings of the Society for Experimental Biology and Medicine*, *224*, 211-215.

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

Pabst, S., Brand, M., & Wolf, O. T. (2013). Stress effects on framed decisions: There are differences for gains and losses. *Frontiers in Behavioral Neuroscience*, *7*, Article 142.

Steneck, N. H. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics*, *12*, 53-74.

Streiner, D. L., & Norman, G. R. (2008). *Health Measurement Scales: A practical guide to their development and use*. Oxford: Oxford University Press.

Tijdink, J. K., Smulders, Y. M., Vergouwen, A. C., de Vet, H. C., & Knol, D. L. (2014). The assessment of publication pressure in medical science; validity and reliability of a Publication Pressure Questionnaire (PPQ). *Quality of Life Research*, *23*, 2055-2062.

Tijdink, J. K., Vergouwen, A. C., & Smulders, Y. M. (2013). Publication pressure and burn out among Dutch medical professors: A nationwide survey. *PLoS ONE*, *8*, Article e73381.

Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions the impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, *60*, 275-304.

Vinkers, C. H., Zorn, J. V., Cornelisse, S., Koot, S., Houtepen, L. C., Olivier, B., . . .Joëls, M. (2013). Time-dependent changes in altruistic punishment following stress. *Psychoneuroendocrinology*, *38*, 1467-1475.

Wells, J. E. (1998). Making good use of numerical predictors: Alternatives to categories or straight lines. *International Journal of Methods in Psychiatric Research*, *7*(1), 66-71.

West, C. P., Tan, A. D., Habermann, T. M., Sloan, J. A., & Shanafelt, T. D. (2009). Association of resident fatigue and distress with perceived medical errors. *Journal of the American Medical Association*, *302*, 1294-1300.

Wicherts, J. M. (2011). Psychology must learn a lesson from fraud case. *Nature*, *480*, 7.

Author Biographies

Joeri K. Tijdink is a psychiatrist working in a community hospital, Tergooi, in Hilversum. He holds a research position at the VU Medical Center in Amsterdam. He started his PhD trajectory on publication pressure and scientific misconduct in 2012, and this article is part of his PhD work.

Reinout Verbeke is a journalist, formerly working at *EOS* magazine (a popular scientific magazine in Flanders) where he conducted this research together with Joeri Tijdink and published epidemiological data of this survey in *EOS*. Currently, he is working as a scientific advisor of the museum of natural sciences in Brussels, Belgium.

Yvo Smulders is a professor of internal medicine at VU medical center in Amsterdam. His research interests include flaws of evidence-based medicine; he is a protagonist of responsible research. He was the primary supervisor of the PhD thesis of Joeri Tijdink.