Publish & Perish

Research on research and researchers

Joeri K. Tijdink

Publish & Perish

Research on research and researchers

colofon

ISBN: 978-90-825123-4-2

opmaak: proefschrift-aio.nl
druk: dpp.nl

VRIJE UNIVERSITEIT

Publish & Perish

Research on research and researchers

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. V. Subramaniam,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Geneeskunde
op donderdag 30 juni 2016 om 9.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Joeri Kees Tijdink
geboren te Tiel

promotoren: prof.dr. Y.M. Smulders
prof.dr. L.M. Bouter

CONTENT

1

Chapter 1

General Introduction

What is this thesis about?

Why should we care?

Publication culture; some historical background

Why do we publish?

Competition

The biomedical scientist

Unresolved issues

Thesis objective

Outline of this thesis

What is this thesis about?

This thesis deals with organizational, cultural and personal factors involved in the execution of biomedical research. More specifically, the prime goal of this thesis is to describe publication culture and to unravel different factors related to biomedical science and scientists that might influence the publication process. I focus on factors related to the scientist him- or herself; perceived publication pressure, burnout symptoms in scientists, self-reported research misbehavior, biases in interpreting scientific information, and personality traits. This has brought me to the title of this thesis. Publish OR Perish has always been and still is a well-known adagium by academic leaders (1;2). Due to the results of this thesis, I have slightly changed this saying; our results point out that you can publish your results and still perish in present culture due to publication pressure and emotional exhaustion.

Why should we care?

Publishing scientific articles is the core business of academic researchers. Together, Dutch scientists deliver more than 30.000 peer reviewed papers each year. (3) Due to heavy workload and other stressful aspects of scientists' professional life, the numerous steps towards answering a research question are susceptible to bias and intentional and unintentional errors. This can jeopardize the validity and the credibility of the scientific enterprise. Furthermore, most scientists rely on earlier evidence collected by others to design their experiments and line of reasoning. If this foundation is invalid, new knowledge is prone to corruption by the biases of the past. Biomedical sciences in particular have a vulnerable position, as they provide results that may directly influence patients' health. So, yes, we should care.

Publication culture; some historical background

To accurately understand current publication culture, we should look at the history of the process of biomedical science. Although the first scientific manuscripts date back from 400 BCE (Hippocrates) and the first 'peer reviewed process' is described in a book called Ethics of the Physician (900 AD) (4), the present form of publishing scientific articles through peer review as standard for defining scientific quality is fairly 'new'. It was around 1940 that JAMA started using experts ('peers') to review scientific articles (5). With the expansion of information technology and the Internet, publishing scientific results is currently transforming rapidly. The number of publications has increased dramatically in the last decade and every day new, digital open access scientific journals pop up (6). This has made scientific publishing a fast, streamlined, enterprise, partly driven by financial incentives.

Why do we publish?

On the one hand, publishing gives researchers the opportunity to present the results of their studies, compare findings with the body of evidence in the field, and convey a personal line of reasoning to the scientific community. Scientists are part of the knowledge circle; adding valid information to the existing body of knowledge that will bring the research field at issue to the next level. Furthermore, publications are career cornerstones. Without publishing scientific work, nobody will recognize you as a scientist or expert. Publications are essential for an academic career, particularly if they are published in high-impact journals. Finally, it is rewarding to present your results to the scientific community; you receive recognition and status.

The tacit reward system was formed earlier in the 20^{th} century, following a period in which education, research and public's interest were assessed in an informal manner, and scientific work was prioritized by policy makers and scholars (7). This reward system was replaced by a more 'measurable' and individual-oriented reward system in the eighties of the 20^{th} century. This change in allocating of funds and rewards has provoked a more 'market-like' competition among scientists (8).

Competition has also changed the evaluation process; productivity in terms of publications and citations is nowadays the dominant pillars of the assessment of individuals and organizations (9). One of the consequences is that individual productivity and aggregate output have increased enormously in the past decades. However, if a reward system is not functioning properly, it can backfire. Behavioral research has shown us that one-sided reward systems cause problems (10); some people will mainly focus on the measurable parameters and pay less attention to content.

Competition

Competition is often seen as a salutary driving force for the scientific enterprise; it creates an atmosphere of high performance and can stimulate researchers to improve their rating and the rating of their institute.

Many believe that these beneficial consequences dominate over the negative effects; people get the best out of themselves by competing with others and push scientific fields to their limits to be the first to report a discovery (7). Competition provides incentives to individuals to excel. Another perceived advantage of competition is that rivalries can provide a counterweight to confirmation bias – the tendency to favor evidence that supports dominant beliefs. Scientists will be eager to disprove, surpass or scoop discoveries made by rivals.

Science has thus adopted a type of 'tournament structure', with a steep drop off in reward terms for those who don't come in first: the winner takes all (11). Competition is clearly essential in individual sports, pushing yourself to your (physical) limit. But is this also be the case for science? To what extent should science provoke and reward individual glory? Or should science be a team sport, with collaboration and cooperation as its fundamental values and where scientific performance is rated by team results rather than by individual indicators?

Several well-known scientists within different disciplines have discussed and criticized contemporary publishing culture ('publish or perish') (1;8;12-14). As mentioned earlier, myopic rewarding of scientific output ignores other important academic duties of a scientist such as education, mentoring and collaboration.

The holy grail for scientists of today's competition is publishing in high impact journals. (2) It will further ones career and foster individual status and prestige. This will be easiest for confirmatory results, or at least results that confirm mainstream thinking, and indeed these positive results are not only more frequently published, the negative ones are increasingly disregarded and less likely to be published. The result of this is referred to as 'publication bias'. Factors such as publication pressure and publication bias may compromise objectivity and integrity of research, as they drive scientists towards production of easily publishable results at all costs. (15)To quote Ioannidis (author of the classical paper 'why most research findings are false' {Ioannidis, 2005 562 /id}) 'competition and conflicts of interest distort too many medical findings' {Ioannidis, 2011 528 /id}

Furthermore, a 'winner takes all competition' may provoke fraud or other research misbehavior (18). It can lead to sabotage of competitors'work, biased peer review and engagement in other Questionable Research Practices (QRP) or more severe research misbehavior like data fabrication (19). A relationship between level of competition and research misbehavior has indeed been suggested in earlier research. (20)

An essential aspect of competition is competing for funds. Due to smaller budgets of funding institutions and increasing number of scientists, less and less funding is available per scientists. Success rates of less than 10% are no exception, and funding proposals are increasingly judged by the publication record of the applicant (11).

Strong competition can compromise collaboration. Fear of being scooped (21) may bring researchers to share less information, methods and materials, and work in secrecy (22). All these aspects might contribute to a hypercompetitive culture focused on scientific output rather than on scientific quality. These detrimental aspects may counteract the benefits of competitiveness. There is a paucity of empirical evidence on these aspects of the publication culture. The evidence that does exist is mostly anecdotal or extracted from qualitative research, and pictures a culture of hypercompetition (13), where scientists are reported to 'rush into print, cut corners, exaggerate their findings and overstate the importance of their results' (23;24). These processes may have resulted in a large amount of research waste (24;25).

The biomedical scientist

Biomedical scientists are the heart of the biomedical scientific community and are the study objects in this thesis. How do biomedical scientists work? What are their drivers? How do they think and feel about their profession?

Scientists are often seen as highly engaged and honest persons, who are mainly rational in their thinking. But are they truly rational and honest when they are facing conflict-of-interest decisions? Can they resist the temptation to follow selfish incentives at the expense of their moral beliefs and viewpoints? Or can they cut corners without feeling bad about themselves? (26) Can they indeed downplay their emotional states in order to pursue highly regarded moral behavior? In other words, are scientists predominantly 'homo sapiens' (= able to know), or 'homo sentiens' (= able to feel)?

Emotional states may influence scientific practice substantially, but are rarely considered in discussions addressing flaws in science. A biomedical scientist alone in a room, behind a computer is typically aware of other pressing academic tasks such as patient care and education, of managers demanding scientific output, and of the need for continued funding. Added to those are temporary contracts, envious colleagues, and career psychopaths who may be around in the very same department, seeking victims.

Moreover, they have so many things to bear in mind. With all the aforementioned factors it is almost inevitable that biomedical scientists are not always able to produce flawless results. They are continuously subject to their own experiences and feelings. Their thoughts and decisions are likely to be influenced by such factors (27;28). This thesis tries to shed some light on these understudied aspects of publication culture.

Unresolved issues

To start with, there is some evidence that biomedical scientists are profoundly stressed and that this might be caused by persistent institutional demands to publish (preferably in high impact journals) (29). Is the biomedical scientist able to withstand these demands? Or does the driving force to publish originate in the individuals themselves? Should we consider the pressure scientists experience and the emotional states it evokes , as an important risk factor, possibly resulting in mental problems? And could this pressure be (partially) accountable for questionable research practices and even lead to research misconduct?

In recent years, several cases of severe violation of scientific integrity (such as the Stapel and Poldermans cases) (30-32) caused great concern in the Netherlands. These cases have led to an intense debate about how widespread fraudulent behavior is, and what its causes and consequences are. These concerns are shared by many scientists and academic leaders at university medical centers and have resulted in more awareness of the threat of scientific misconduct. This has led to a number of actions – typically in the form of stricter rules and obligatory education - , but unfortunately the evidence base for these measures is poor.

Breaches of scientific integrity involve a wide spectrum of behavior, ranging from outright fraud, such as data fabrication and falsification, to much more common, but less well known, questionable research practices (QRP). Examples of QRPs are adding an author to your authors list without significant contribution or rounding of a p-value incorrectly. It is unknown to what extent QRPs are common in the biomedical research, whether they are related to the perceived publication pressure, and how biomedical scientists asses QRPs in terms of prevalence, causes and consequences.

In addition, it can be questioned which – if any - psychological characteristics influence the integrity of scientific practice? In other words, can specific personality traits be identified that are associated with research misbehavior, and do these personality profiles and corresponding misbehavior practices differ between different hierarchical positions in academia?

Thesis objective

This thesis aims to identify major personal factors that influence the process of publishing. Special attention will be given to perceived publication pressure. This thesis attempts to unravel the effects of contemporary publication culture on professional behavior, mental problems and personality traits of scientists, and vice versa. It will also focus on the effects of positive outcome bias and research funding on the credibility of publications.

Outline of this Thesis

As in most dissertations in biomedicine, the chapters in this thesis are based on different scientific publications. As a result, some chapters may have some similarities due to the use of the same data and equal methodology. In this thesis, chapter 2,3 and 4 are based on the same dataset and chapter 8 and 9 are based on the same dataset.

Chapter 1 is the general introduction in this thesis. **Chapter 2** investigates the reliability and validity of a publication pressure questionnaire. **Chapter 3 and 4** uses the same study sample as chapter 2. **Chapter 3** addresses emotional exhaustion in biomedical professors. **Chapter 4** is our first exploratory study to investigate publication pressure among biomedical professors in the Netherlands. **Chapter 5** investigates whether publication pressure is correlated to research misbehavior among researchers in different academic ranks. **Chapter 6** explores the perceptions of contemporary publication culture in a qualitative approach by reporting on the results of 12 focus group discussions among biomedical researchers in 4 University Medical Centers in the Netherlands. **Chapter 7** studies the influence of pharmaceutical funding and reported study outcome on perceived credibility of a scientific abstract in a sample of Dutch psychiatrists. **Chapter 8** investigates the role of specific personality traits that might be associated with questionable research practices, in relation to academic hierarchy.

Chapter 9 should be interpreted differently. It is meant as a cheerful 'tongue in cheek' article that is based on the real data from chapter 8 but twisted for this special purpose[1]. It describes a cluster analysis and determines a certain personality profile of successful scientists that engages is research misbehavior. Finally, **Chapter 10** rounds off this thesis by integrating the highlighted findings, discussing their relevance and offering an outlook for future projects.

---

1 The text was written for the BMJ Christmas issue, but not accepted. It's currently under review by another journal.

Reference List

(1) Publish or perish. Nature 2015;521;159.

(2) Miller AN, Taylor SG, Bedeian AG. Publish or perish: academic life as management faculty live it. Career development international 2011;16:422-45.

(3) Chiong Meza C., Van Steen J, de Jonge J. De Nederlandse universitair medische centra. Den Haag: Rathenau Instituut; 2014.

(4) Spier R. The history of the peer-review process. Trends Biotechnol 2002;20:357-8.

(5) Burnham JC. The evolution of editorial peer review. JAMA 1990;263:1323-9.

(6) Ridker PM, Rifai N. Expanding options for scientific publication: is more always better? Circulation 2013;127:155-6.

(7) Merton R.K. Priorities in scientific discovery. Am Sociol Rev 1957;22:635-59.

(8) van Dalen HP, Henkens K. Intended and unintended consequences of a publish-or-perish culture: A worldwide survey. Journal of the American Society for Information Science and Technology 2012;63:1282-93.

(9) Walker RL, Sykes LF, Hemmelgarn BR, Quan H. Authors' opinions on publication in relation to annual performance assessment. BMC Medical Education 2010; 10:21.

(10) Ariely D, Gneezy U, Loewenstein G, Mazar N. Large Stakes and Big Mistakes. The Review of Economic Studies 2009;76:451-69.

(11) Fang FC, Casadevall A. Competitive Science: Is Competition Ruining Science? Infect Immun 2015;83:1229-33.

(12) Adler N, Harzing AW. When Knowledge Wins: Transcending the Sense and Nonsense of Academic Rankings. Academy of Management Learning and Education 2009;8:72-95.

(13) Anderson MS, Ronning EA, DeVries R., Martinson BC. The perverse effects of competition on scientists' work and relationships. Science and Engineering Ethics 2007;13:437-61.

(14) Fanelli D. Do pressures to publish increase scientists' bias? An empirical support from US States Data. PLoS One 2010;5:e10271.

(15) Fanelli D. Negative results are disappearing from most disciplines and countries. Scientometrics 2012;90:891-904.

(16) Ioannidis JP. Why most published research findings are false. PLoS One 2005;2:e124.

(17) Ioannidis JP. An epidemic of false claims. Competition and conflicts of interest distort too many medical findings. Sci Am 2011; 304:16

(18) Bedeian AG, Van Fleet DD, Hyman HH. Scientific achievement and editorial-board membership. Organizational Research Methods 2009;12:238.

(19) Anderson MS, Ronning EA, De VR, Martinson BC. The perverse effects of competition on scientists' work and relationships. Sci Eng Ethics 2007;13:437-61.

(20) Combating scientific misconduct. Nat Cell Biol 2011;13:1.

(21) Sullivan D. Competition in bio-medical science: Extent, structure, and consequences. Sociology of Education 1975;48:223-41.
(22) McCain K.W. Communication, competition and secrecy: The production and dissemination of research-related information in genetics. Sci Technol Hum Val 1991;16:491-516.
(23) Alberts B, Kirschner MW, Tilghman S, Varmus H. Rescuing US biomedical research from its systemic flaws. Proc Natl Acad Sci U S A 2014;111:5773-7.
(24) Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet 2014;383:166-75.
(25) Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP, et al. Biomedical research: increasing value, reducing waste. Lancet 2014;383:101-4.
(26) Mazar NF, Ariely D. Dishonesty in scientific research. J Clin Invest 2015;125:3993-6.
(27) Ariely D. Predictably Irrational: The Hidden Forces That Shape Our Decisions. HarperCollins; 2008.
(28) Kahnemann D. Thinking Fast and Slow. Penguin Books Ltd; 2011.
(29) Miller AN, Taylor SG, Bedeian AG. Publish or perish: academic life as management faculty live it. Career development international 2011;16:422-45.
(30) Levelt Committee, Noort Committee, Drenth Committee. Flawed science: The fraudulent research practices of social psychologist Diederik Stapel. 2012.
(31) Onderzoekscommissie wetenschappelijke integriteit. Rapport vervolgonderzoek naar mogelijke schending van de wetenschappelijke integriteit 2012. 2012.
(32) Wicherts JM. Psychology must learn a lesson from fraud case. Nature 2011;480:7.

2

Chapter 2

2

The assessment of publication pressure in medical science; validity and reliability of a Publication Pressure Questionnaire (PPQ)

J.K. Tijdink, Y.M. Smulders, A.C.M. Vergouwen, H.C.W. de Vet, D. L. Knol

Qual Life Res 2014;23;2055-2062, doi: 10.1007/s11136-014-0643-6

Abstract

Purpose

To determine content validity, structural validity, construct validity and reliability of an internet-based questionnaire designed for assessment of publication pressure experienced by medical scientists.

Methods

The Publication Pressure Questionnaire (PPQ) was designed to assess psychological pressure to publish scientific papers. Content validity was evaluated by collecting independent comments from external experts (n = 7) on the construct, comprehensiveness and relevance of the PPQ. Structural validity was assessed by factor analysis and item response theory (IRT) using the generalized partial credit model. Pearson's correlation coefficients were calculated to assess potential correlations with the emotional exhaustion and depersonalization subscales of the Maslach Burnout Inventory (MBI). Single test reliability (lambda2) was obtained from the IRT analysis.

Results

Content validity was satisfactory. Confirmatory factor analysis did not support the presence of three initially assumed separate domains of publication pressure (i.e., personally experienced publication pressure, publication pressure in general, pressure on position of scientist). After exclusion of the third domain (six items), we performed exploratory factor analysis and IRT. The goodness-of-fit statistics for the IRT assuming a single dimension were satisfactory when four items were removed, resulting in 14 items of the final PPQ. Correlations with the emotional exhaustion and depersonalization scales of the MBI were 0.34 and 0.31, respectively, supporting construct validity. Single test administration reliability lambda2 was 0.69 and 0.90 on the test scores and expected a posteriori scores, respectively.

Conclusion

The PPQ seems a valid and reliable instrument to measure publication pressure among medical scientists.

Introduction

Scientific output is increasingly being used as a performance parameter for both academic medical centres as well as individual medical scientists (1;2). Quantitative measures of scientific performance, such as the Hirsch index (3), have become particularly important, as these influence not only fund raising potential, but also financial rewards (such as salary and grants allocations), career development possibilities, and prestige and status (1;4-6). This increased emphasis on scientific output is accompanied by an enormous increase in publications in the past decades. For example, the number of scientific journals increased from 5000 in 1997 to 8000 journals in 2010 (ISI Web of Knowledge, http://www.webofknowledge.com, consulted March 2013), and the amount of scientific papers doubles every 12 years (http://www.scopus.com).

The importance of proxy measures of scientific performance such as the journal impact factor or the Hirsch index, and the emphasis on quantitative aspects thereof, increase pressure on medical professionals to publish, and intensifies competition between them. Although this competition for papers and funding is often considered a salutary driving force among scientists, increasing efficiency and productivity(7), potential negative effects are less often highlighted (8;9). Concerns have been expressed that scientists are continuously producing 'publishable' results at the expense of quality, validity, scientific rigour and personal integrity. As a result, clinical practice based on research outcomes can be jeopardised (10;11). Since stress and burnout symptoms are common among medical doctors and residents (12;13), pressure to increase the amount of scientific publications may also affect their mental well-being. This could in turn affect their professional performance, including patient care, education, and research (9;14)H.

There is no available measurement instrument to assess the degree of publication pressure that scientists experience. Such an instrument could be of value for further research on causes and consequences of publication pressure such as the influence of publication pressure on the quality of the publications and on mental health and performance of scientists.

We developed a brief questionnaire to assess publication pressure. The questionnaire was tested and distributed among all Dutch medical professors. Based on these data, the aim of the current study is to assess the content validity, structural validity, construct validity and reliability of the Publication Pressure Questionnaire, and to further improve it.

Methods

Publication Pressure Questionnaire

We designed a 24-item Publication Pressure Questionnaire (PPQ) to assess pressure among medical scientists to publish scientific papers. We hypothesized a theory of publication pressure based on a relation between increasing demands by policy makers on publication output, the increased amount of published scientific papers, and perceived increased competition among medical scientists. The Publication Pressure Questionnaire (PPQ) is a self-reported measure. The draft version consisted of 24 items with a 5 point response scale (range 0-4, 0 = totally disagree; 4 = totally agree, Appendix A). Examples of statements used in the questionnaire were 'Without publication pressure, my scientific output would be of higher quality', and 'The scientific output criteria set by my university to evaluate my (re)appointment are stimulating'. We alternately formatted the statements in a positive and negative way to avoid 'yeah-saying' (15). Positive statements were reversely scored (i.e. totally disagree = 4 points instead of 0 points).

Participants

All medical professors (N=1206) working at one of the 8 academic medical centres in The Netherlands were sent an invitational e-mail in September 2011. This e-mail explained the objectives of the study and provided them with a link to an anonymous online questionnaire on a protected website. Those who did not respond were sent a reminder after 3 weeks.

Content validity

The test construct was evaluated using expert opinions on contemporary publication culture.

Based on consensus standards for the selection of health measurement instruments (COSMIN), four requirements for good content validity were defined (16):

1. All items should refer to relevant aspects of the construct to be measured.
2. All items should be relevant for the study population (e.g. age, gender, disease characteristics, country, setting).
3. All items should be relevant for the purpose of the measurement instrument (discriminative, evaluative, and/or predictive).
4. All items together should comprehensively reflect the construct to be measured.

To evaluate the first, second and third requirement, a convenience sample of 5 experienced medical scientists in different research fields and 2 mental coaches (specialized in the mental support of medical professors) in the departments of Internal Medicine and Psychiatry in the affiliated hospitals of JT, AV and YS were invited to comment on the construct of the PPQ during the process of its development. Specifically, they reflected on what the PPQ aims to measure, and judged the relevance of the 24 items. After this process, three of the authors (JT, AV and YS) proposed 3 separate domains of publication pressure: (1) personally experiences of publication pressure, (2) publication pressure in general and (3) pressure on the formal job position. The authors of this paper and the aforementioned expert panel judged if all items were relevant for evaluative purposes, and the same group of experts judged the comprehensiveness of the questionnaire.

Structural validity and reliability

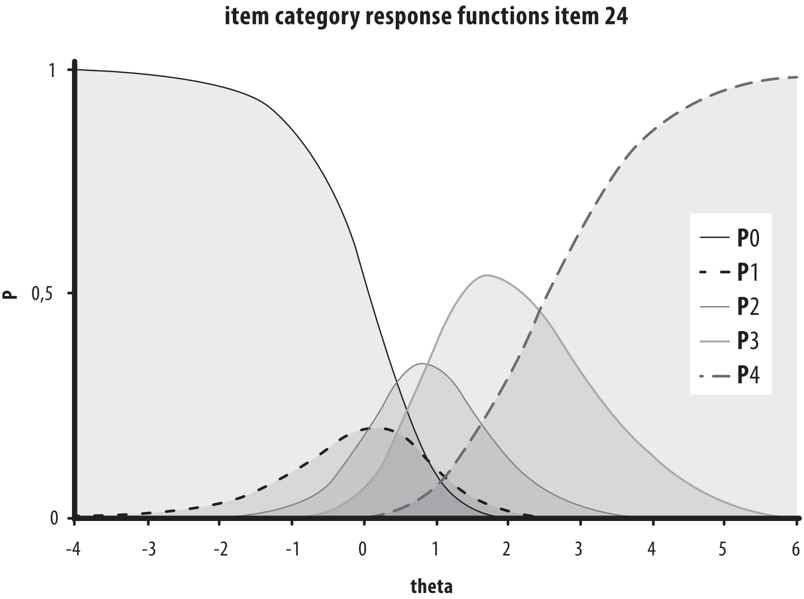
The initial model with 3 domains of publication pressure was evaluated using confirmatory factor analysis (CFA) for ordered categorical items, which uses the matrix of polychoric correlations as input (17). The criteria for a well-fitting model are root mean square error of approximation (RMSEA) <0.06, comparative fit index (CFI) and Tucker-Lewis index (TLI) both >0.95, and weighted root mean square residual (WRMR) <1 or standardized root mean square residual (SRMR) <0.08 (18) . If the proposed model did not fit within the CFA, exploratory factor analysis (EFA) was carried out to assess the dimensionality of the item pool. All factor analyses were conducted in Mplus 6.12 (19), using the estimation method of weighted least squares with mean and variance adjustment (WLSMV) (18).

The fit (of each subscale) was then further investigated by means of item response theory (IRT). In IRT models, the response probabilities of each person to the individual items are modelled as a function of the latent trait (θ) of that person. This makes IRT models particularly well-suited for analyzing item fit and differential item functioning (20). One of the most flexible IRT models is the generalized partial credit model (GPCM), which was used in the sequel (21). In the GPCM, the probability of score *j* on item *i* as a function of the latent trait θ of a person is given by the item category response functions (e.g., item 24 in Figure 1)

$$P_{ij}(\theta) = \frac{\exp[\alpha_i(j\theta - \sum_{g=1}^{j}\beta_{ig})]}{1 + \sum_{h=1}^{m_i}\exp[\alpha_i(h\theta - \sum_{g=1}^{h}\beta_{ig})]} \quad j = 0, \ldots, m_i$$

where m_i + 1 denotes the number of item categories and β_{ij} and α_i are item parameters. The parameter β_{ij} is a category intersection parameter of item *i*, i.e. the point in which the probability of responding in category *j* – 1 is equal to the probability of responding in category *j*. For example, in Figure 1 (item 24) it can be seen that the item category response functions of categories 0 and 1, and of 1 and 2 cross each other at $\theta = \beta_{24,1} = 0.864$ and $\theta = \beta_{24,2} = 0.056$, respectively (see Table 2). It should be noted that strict ordering of the category intersection parameters is not implied by the GPCM, i.e. the category intersection parameters cannot strictly be interpreted as item difficulty parameters. Here, i.e. for item 24, it means that category 1 is for none of the theta-values the modal response category. Finally, α_i is the discrimination parameter that indicates the extent to which the item response is related to the latent scale. Values of α_i greater than 0.64 are considered to reflect a moderate discrimination parameter (22). Item fit was based upon the Lagrange multiplier (LM) item goodness-of-fit tests (23)

Figure 1



The item category response functions of item 24.

where, in a stepwise fashion, an item was removed from the item pool if both the LM test was significant ($P < 0.05$) and the effect size (absolute difference between observed and predicted item scores) was larger than 0.12 (24). (Van Groen et al. (26) considered an effect size of 0.10 indicative of acceptable model violation for a 4-point scale. Therefore, we used an effect size of 0.12 for our 5-point scale.) All GPCM analyses were carried out using the program MIRT (25). Item parameters were estimated using the method of marginal maximum likelihood. Single test-administration reliability coefficient lambda2 (26) of the test score and the reliability of the estimated trait scores were computed. Trait scores were estimated as expected a posteriori (EAP) scores (27).

Construct validity

To prepare the perfect conditions for an ideal constuct validation ideally a gold standard is needed that should be used as a reference for the new scale that is being validated. Since such scale did not exist, we used the MBI to have a validated measure that can be used as a reference point. Hypothesis testing, assisting in assessing construct validity, was conducted by comparing the scores on the PPQ to the scores of the translated version of the Maslach Burnout Inventory (28;29), an instrument designed as a measure of work related stress.

The MBI consists of 20 items divided in three domains; emotional exhaustion (8 items, key feature), depersonalisation (5 items) and personal accomplishment (7 items). Emotional exhaustion is characterized by total loss of energy at work and a negative attitude towards work-related activities. Depersonalisation is a form of alienation from work, where a person has lost interest in his work and his colleagues. Personal accomplishment is a positive symptom as it is the feeling of capability and satisfaction in doing work.

A moderate correlation (ranging 0.3-0.5) between the PPQ (raw sum score and EAP scores) and the two burn out domains on the MBI (emotion exhaustion and depersonalisation) was expected because of partial overlap between the negative symptoms of burn out (emotional exhaustion and depersonalisation) and publication pressure, as both closely relate to work-related stress. We did not expect overlap with personal accomplishment since this construct is not directly related to publication pressure.

Results

In total, 578 (48%) medical professors responded, of whom 437 (36%) completed the questionnaire. The demographics of the participants are tabulated in table 1.

Content validity

The experts and coaches reflected on the PPQ items and on what they actually measure. Besides the general construct of pressure in producing and publishing scientific papers, additional constructs were mentioned such as personally

Table 1. Demographics

Demographics			
		N=437	%
Gender	Male	345	79
	Female	92	21
Age	26-45	36	8
	46-55	206	47
	56 and older	195	45
Marital status	Married or cohabiting	401	92
	Single	36	8
Home living children	None	217	50
	1	56	13
	2	96	22
	3 or more	68	15
Years of professorship	0-5	150	34
	6-10	129	30
	11-15	86	20
	15 or more	72	16
Nr. 1 Work priority	Research	255	59
	Education	40	9
	Patient care	63	14
	Management	79	18

experienced publication pressure, publication pressure in general, shortcomings in publication culture, and position-related problems.

Based on this feedback, it was concluded that the PPQ probably measures more than just publication pressure. The term 'publication culture' was then introduced in the invitational e-mail, which had the additional advantage of being a more neutral term, causing less risk for selective non-response and framing.

Table 2. Estimated item discrimination (αi) and category intersection (βij) parameters of the final GPCM consisting of 14 items.

Item	αi	βi1	βi2	βi3	βi4
1	0.767	-0.868	0.194	0.858	1.322
3R*	0.255	-1.299	1.115	2.175	1.466
6	0.715	-0.697	0.196	1.044	2.217
7R	0.546	-2.261	-0.766	0.702	0.673
8	0.758	-1.110	0.217	0.187	1.658
10	0.704	-1.300	-0.634	0.465	1.875
11R	0.594	-2.664	-1.660	-0.149	0.753
12	0.914	-2.357	-0.703	0.356	1.840
13	1.328	-2.679	-0.739	-0.532	1.831
15R	0.250	-2.172	0.220	0.319	1.099
17	0.263	-1.185	0.392	0.427	1.610
19R	0.443	-2.845	-0.275	1.038	1.399
21	0.397	-0.224	0.720	0.923	2.130
24	1.188	0.864	0.056	0.863	2.439

R = positive worded item scores reversed; theta standard normal

The experts considered the 24 items to be relevant (COSMIN criteria 1-3) and comprehensive (COSMIN 4). There were no suggestions for additional/new items, but the experts did suggest to rephrase some items.

Structural validity and reliability

The data from 437 respondents (78.8 % male, age ranging from 30-75) were used for statistical analysis. Using CFA, the initial 3-factor model did not fit well: the goodness of fit statistics were: RMSEA = 0.125, CFI = 0.663, TLI = 0.617, WRMR = 2.560.

In particular, subscale 3 (pressure on the subjects position) was not recognized as a separate domain. Based on the content of these 6 items, which did not

Table 3. Lagrange multiplier item goodness of fit tests for PPQ of 14 items.

Item	Lagrange multiplier*	p-value	Absolute Difference
1	0.03	.99	0.00
3R**	4.26	.12	0.06
6	1.62	.45	0.02
7R	2.49	.29	0.03
8	6.56	.04	0.09
10	2.45	.29	0.05
11R	2.04	.36	0.05
12	4.08	.13	0.03
13	15.60	.00	0.05
15R	8.48	.01	0.07
17	11.00	.00	0.11
19R	3.42	.18	0.06
21	0.11	.95	0.01
24	13.24	.00	0.05

* Chi-squared with df = 2; ** R = positive worded item scores reversed

reflect publication pressure, but more its consequences, we excluded these items, and combined the remaining 2 subscales (personally experienced publication pressure, and publication pressure in general). Although the resulting unidimensional model (18 items) did not fit well (RMSEA = 0.128, CFI = 0.710, TLI = 0.672, SRMR = 0.113), EFA revealed no clear multidimensional structure, and it was decided to further improve the model fit using IRT, i.e. the GPCM.

Starting with the remaining 18 items, items were removed from the item pool in a stepwise fashion based on the LM item goodness-of-fit tests. After removing items 2, 18, 5 and 9, respectively, all absolute differences were <0.12. Table 2 gives the item discrimination and category intersection parameter estimates of the final 14 items of the PPQ, and table 3 the LM statistics. Single test administration reliability lambda2 of the test scores and the EAP scores of the PPQ were 0.69 and 0.90, respectively.

Construct validity – hypothesis testing

The mean sum score of the 14 item PPQ was 25.8 (SD 6.1, range 0 to 43). Pearson correlations of 0.34 (95% CI 0.25 – 0.42) and 0.31 (95% CI 0.22 – 0.39) were found between the two selected domains (emotional exhaustion and depersonalisation) of the MBI and the 14 item PPQ's sum score, indicating moderate correlation. Using the EAP scores of the PPQ, the correlation with the emotional exhaustion scale increased (0.45: 95% CI 0.37 – 0.52), whereas it remained similar for the depersonalisation scale (0.29: 95% CI 0.20 – 0.37)

Discussion

In a large sample of medical professors, the draft version of the PPQ was tested using up-to-date methods of questionnaire development and evaluation. To our knowledge, this is the first evaluation and validation of a questionnaire measuring publication pressure among medical scientists. Regarding the content validity, the four COSMIN criteria were satisfied. The questionnaire appears to be reliable (0.90), but it has relatively low discrimination parameters (see table 2). As expected, the PPQ correlated moderately with the subscales emotional exhaustion and depersonalisation of the Maslach Burnout Inventory. This correlation supports construct validity of the questionnaire, although further assessment of construct validation is desirable.

Interpretation of results

The questionnaire was originally developed as a three dimensional questionnaire. Using CFA it appeared that the a-priori hypothesis of 3 domains could not be confirmed, and the third domain (addressing consequences of pressure on the job position) was dropped. These items rather refer to the consequences of publication pressure than to publication pressure itself. Consequences of publication pressure itself may be a topic for further research. Thus it was excluded as part of the measurement instrument. After IRT a unidimensional scale was defendable. Although this scale had a high reliability of the EAP scores (0.90), lambda2 reliability (0.69) was just below 0.70, which is frequently used as a minimal reliability. Secondly, the estimated item discrimination parameters of the GPCM were moderate, with only seven of the 14 item discrimination parameters (see table 2) of 0.64 or higher (22). Most positive stated items (reverse scored) had low discrimination parameters. Possibly this is due to the statements which were formulated more neutrally such as item 3 'My scientific publications contribute to better (future) medical care'.

The descriptives of total PPQ score (mean 25.8, SD 6.1 range 0 – 43) show that medical professors do not vary a lot on total PPQ score. This homogeneity of the study population with respect to the experience publication pressure may be an important explanation for the low to moderate discrimination parameters.

Thirdly, construct validity was assessed by correlating the PPQ with the MBI (a questionnaire measuring burn out symptoms). The summed PPQ score correlated moderately (0.34) with the MBI and measures a slightly different construct. Burn out is a much broader construct than publication pressure alone, and the MBI is designed for all labour people, not just publishing scientists, resulting in an expected moderate correlation.

Strength and limitations

An important strength is the large sample size of 437 professors completing the questionnaire. Although the response rate compares favourably to response rates in similar types of surveys (30) and might be selective, this might affect the results of the experienced level of publication pressure by medical professors.

We think that the response bias in our study may have bidirectional; non-response may be related to lack of time or sense of task overload. This has caused underestimation of pressure and of discontent with publication pressure.

On the other hand, medical professors who don't have publication pressure may consider publication pressure irrelevant, and thus decline participation. Such bias would conceivably have caused overestimation of pressure and disapproval of publication pressure among respondents.

However, this response bias is important in the estimation of the prevalence of publication pressure among medical professors, but less so in our study on the psychometric analysis of the questionnaire. In the assessment of construct validity there is across sectional comparison of scores on the PPQ and other instruments. Non-response has only influence if it highly affects the range of values of publication pressure. This is not to be expected. Also content validity and structural validity are hardly or not affected by selective non-response.

A limitation of our study is that, for construct validity, we compared the PPQ with only one instrument; the MBI. The MBI is regularly used in similar types of studies addressing work-related psychological stress in medical professionals. Burn-out is not a specific, but in our view a conceivable consequence of publication pressure, which is indeed supported by the literature (9;12).

Although we found the anticipated correlations with the MBI subscales, the use of more than one comparing measurement improves assessment of the construct validity.

Some other limitations should be acknowledged. Since the PPQ is a questionnaire in Dutch, it was formally translated by an official translation office in English for this manuscript (Appendix A). Besides that, it must be taken into account that the way in which the amount of publication of scientific articles contributes to prestige and success of individual scientists in the Netherlands might differ from other countries. Although this precludes generalisation of the results, publication pressure is reported to be a problem in other Western countries as well (9).

Another issue is the period in which the questionnaire was developed and sent out. During this period, recent fraud cases in the Netherlands were discovered (31), and an intense media and societal debate on scientific integrity sparked, implicating publication pressure as a possible cause for fraudulent behaviour. This could have influenced test results in both directions: aggravation of the pressure on one side, or reluctance to admit pressure on the other side.

The questionnaire is designed to measure publication pressure among medical scientists. However, the questionnaire has only been tested in a specific population (medical professors), and should therefore be tested among other professionals involved in medical research (i.e. associate and assistant professors, post doc fellows and PhD students) and in other research fields (i.e. social sciences or mathematics) as well.

First results of determinants of publication pressure show that years of professorship are related with a higher level of publication pressure (32) Marital status, age and gender of the respondents did not differ significantly in total score of the PPQ questionnaire (32). Furthermore the study population did not differ from the total number of professors as for example gender (http://www.stichtingdebeauvoir.nl/wp-content/uploads/Monitor_Vrouwelijke_Hoogleraren_2012.pdf). Future research could point out relations with questionable research practices, psychological stress or personality traits.

Conclusion

The PPQ is a suitable instrument to measure publication pressure among medical scientist, with acceptable psychometric properties. The assessment of publication pressure enables research into the determinants and consequences of publication pressure, and contributes to the current debate on scientific publications. Further research is needed to evaluate its psychometric properties in other relevant research populations such as PhD students, postdoc fellows and assistant professors.

Reference List

(1) Gannon F. the impact of the impact factor. EMBO Rep 2000;1:293.

(2) Retzer V, Jurasinski G. Towards objectivity in research evaluation using bibliometric indicators- a protocol for incorporating complexity. Basic and Applied Ecology 2009;10:393-400.

(3) Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci U S A 2005;102:16569-72.

(4) Bedeian AG, Van Fleet DD, Hyman HH. Scientific achievement and editorial-board membership. Organizational Research Methods 2009;12:238.

(5) Bird SJ. Research ethics, research integrity and the responsible conduct of research. Science and Engineering Ethics 2006;12:412.

(6) Glick WH, Miller CC, Cardinal LB. Making a life in the field of organization science. Journal of Organizational Behavior 2007;28:817-35.

(7) Anderson MS, Ronning EA, De Vries R., Martinson BC. The perverse effects of competition on scientists' work and relationships. Science and Engineering Ethics 2007;13:437-61.

(8) Anderson MS, Martinson BC, De Vries R. Normative dissonance in science: results from a notional survey of U.S. Scientists. J Empir Res Hum Res Ethics 2007;2:3-14.

(9) Miller AN, Taylor SG, Bedeian AG. Publish or perish: academic life as management faculty live it. Career development international 2011;16:422-45.

(10) Fanelli D. Do pressures to publish increase scientists' bias? An empirical support from US States Data. PLoS One 2010;5:e10271.

(11) West CP, Tan AD, Habermann TM, Sloan JA, Shanafelt TD. Association of resident fatigue and distress with perceived medical errors. JAMA 2009;302:1294-300.

(12) Prins JT, Hoekstra-Weebers JE, Gazendam-Donofrio SM, Dillingh GS, Bakker AB, Huisman M, et al. Burnout and engagement among resident doctors in the Netherlands: a national study. Med Educ 2010;44:236-47.

(13) Shanafelt TD, Boone S, Tan L, Dyrbye LN, Sotile W, Satele D, et al. Burnout and Satisfaction With Work-Life Balance Among US Physicians Relative to the General US Population. Arch Intern Med 2012;172:1377-85.

(14) van Dalen HP, Henkens K. Intended and unintended consequences of a publish-or-perish culture: A worldwide survey. Journal of the American Society for Information Science and Technology 2012;63:1282-93.

(15) Streiner DL, Norman GR. Health Measurement Scales : A practical guide to their development and use. Oxford University Press; 2008.

(16) Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: an international Delphi study. Qual Life Res 2010;19:539-49.
(17) Wirth RJ, Edwards MC. Item factor analysis: current approaches and future directions. Psychol Methods 2007;12:58-79.
(18) Wirth RJ, Edwards MC. Item factor analysis: Current approaches and future directions. Psychological Methods 2007;12:58-79.
(19) Mplus User's Guide [computer program]. Version 6.12. Los Angeles: Muthén & Muthén; 2007.
(20) Embretson S.E., Reise S.P. Item respons theory for psychologists. 1st edition ed. Mahwah, New Jersey: Lawrence Erlbaum Associates; 2000.
(21) Muraki E. A generalized partial credit model: Application of an EM algorithm. Applied Psychological measurement 1992;16:159-76.
(22) Baker FB. The basics of item response theory. 2nd. ed. College Park: ERIC Clearinghouse on Assessment and Evaluation; 2001.
(23) Glas CAW. Modifications indices for the 2-PL and the nominal response model. Psychometrika 1999;64:273-94.
(24) van Groen MM, ten Klooster PM, Taal E, van de Laar MAFJ, Glas CAW. Application of the health assessment questionnaire disability index to various rheumatic diseases. Quality of Life Research 2010;19:1255-63.
(25) Preliminary Manual of the software program Multidimensional Item Response Theory (MIRT) [computer program]. Version 1.01. Enschede: University of Twente; 2010.
(26) Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika 2009;74:107-20.
(27) Skrondal A, Rabe-Hesketh S. Generalized latend variable modeling: Multilevel, lungitudinal and structural equation models. Boca Raton FL: Chapman & Hall/CRC Press; 2004.
(28) Maslach C, Jackson SE, Leiter MP. Maslach Burnout Inventory Manual. 3rd edn. ed. Palo Alto, CA: Consulting Psychologists Press; 1996.
(29) Schaufeli WB, van Dierendonk D. Utrechtse Burnout Schaal (UBOS), handleiding [Utrecht Burnout Scale, manual]. Utrecht: Swets & Zeitlinger; 2000.
(30) Cook C, Heath F, Thompson RL. A meta-analysis of response rates in web- or internet-based surveys. Educational and Psychological Measurement 2000;60:821-36.
(31) Wicherts JM. Psychology must learn a lesson from fraud case. Nature 2011;480:7.
(32) Tijdink JK, Vergouwen AC, Smulders YM. Publication Pressure and Burn Out among Dutch Medical Professors: A Nationwide Survey. PLoS One 2013;8:e73381.

Appendix A. Publication Pressure Questionnaire

Statement	Specific Domain
1. Without publication pressure, my scientific output would be of higher quality	PPQ1
2. Publication of scientific articles is the most important aspect of my work	PPQ1
3. My scientific publications contribute to better (future) medical care*	PPQ1
4. The number of scientific publications contributes to my status	PPQ3
5. The number of scientific publications by colleagues contributes to their status	PPQ2 and PPQ3
6 I experience my colleagues' assessment of me on the basis of my publications as stressful	PPQ1 and PPQ3
7. I experience the publication criteria formulated by my university for my appointment or re-appointment as professor as a stimulus*	PPQ1 and PPQ3
8. Publication pressure puts pressure on relationships with fellow-researchers	PPQ1 and PPQ2
9. Publication pressure results in me publishing more without it compromising the quality of my scientific work	PPQ1
10. I suspect that publication pressure leads some colleagues (whether intentionally or not) to color data	PPQ2
11. The validity of medical world literature is increased by the publication pressure in scientific centers*	PPQ2
12. Publication pressure leads to serious worldwide doubts about the validity of research results	PPQ2
13. In my opinion the pressure to publish scientific articles has become too high	PPQ1 and PPQ2
14. Fellow-specialists envy me my position of professor	PPQ3
15. The competitive scientific climate stimulates me to publish more*	PPQ1
16. I experience my professorship as a burden and I sometimes long for the time I was not a professor yet	PPQ3
17. My colleagues judge me mainly on the basis of my publications	PPQ1 and PPQ3
18. In spite of the pressure to publish, I enjoy investing in other activities that I feel compliment professorship*	PPQ1
19. Fellow-professors maintain their clinical and teaching skills well, despite publication pressure*	PPQ2
20. Team spirit and collegiality are always decisive in the appointment of a professor at my center*	PPQ3
21. I cannot confide innovative research proposals to my colleagues	PPQ1 and PPQ3
22. I have to spend too much time on management tasks	PPQ3
23. It is difficult to combine being a professor with being a teacher	PPQ3
24. Publication pressure harms science	PPQ2

* reverse scored questions

In bold: statements part of the validated PPQ (further explanation of the questionnaire, see methods section)

PPQ1: pressure to publish personally experienced by the respondent

PPQ2: publication pressure in general terms

PPQ3: publication pressure relating to the scientist's position

3

Chapter 3

Emotional exhaustion and burnout among medical professors; a nationwide survey

3

Joeri K Tijdink, Anton CM Vergouwen, Yvo M Smulders

BMC Medical Education; 2014; 14:183, doi:10.1186/1472-6920-14-183

Abstract

Background

Although job-related burnout and its core feature emotional exhaustion are common among medical professionals and compromise job satisfaction and professional performance, they have never been systematically studied in medical professors, who have central positions in academic medicine.

Methods

We performed an online nationwide survey inviting all 1206 medical professors in The Netherlands to participate. They were asked to fill out the Maslach Burnout Inventory, a 'professional engagement' inventory, and to provide demographic and job-specific data.

Results

A total of 437 Professors completed the questionnaire. Nearly one quarter (23.8%) scored above the cut-off of the definition of emotional exhaustion. Factors related to being in an early career stage (i.e. lower age, fewer years since appointment, having homeliving children, having a relatively low Hirsch index) were significantly associated with higher emotional exhaustion scores. There was a significant inverse correlation between emotional exhaustion and the level of professional engagement.

Conclusions

Early career medical professors have higher scores on emotional exhaustion and may be prone for developing burnout. Based upon this finding, preventive strategies to prevent burnout could be targeted to young professors.

Background

Burnout is described as 'a prolonged response to chronic emotional and interpersonal stressors at work', and is three dimensionally defined by 'emotional exhaustion', 'depersonalisation', and 'reduced personal accomplishment (1). Previous studies suggest that burnout, particularly emotional exhaustion, is common among physicians (2-7), affects morale and productivity, but also reduces quality of care and predisposes to medical errors (8-11).

Reported risk factors for burnout in the general population include being young, single, and childless,. As for job-related factors, home-work interface stress and being at the early stage of a professional career appear to increase risk for burnout. In physicians, risk may be aggravated by job-specific circumstances such as demanding patients, reduced resources, and the threat of liability (12-16).

Opposite of burnout stands engagement, defined as 'a positive, fulfilling, work-related state of mind, characterised by vigour, dedication and absorption' (17), and it has been suggested that strong professional engagement may protect against burnout (18).

Medical professors are in many ways at the heart of the medical community as they act as, educators, managers and, - perhaps most importantly - role models for students, residents and colleagues. However, these same activities and responsibilities may render them vulnerable to job-related stress and burnout.

To our knowledge, there are some studies evaluating burnout symptoms by academic rank (19;20) although these symptoms have never been systematically studied in the unique subgroup of medical professors. This study addresses the prevalence, severity and potential determinants of burnout symptoms among medical professors in The Netherlands. Since emotional exhaustion is the core feature of burnout (1;21-24), the association of emotional exhaustion with personal and job characteristics, with the Hirsch index as a measure of scientific success, and with the level of professional engagement was examined in detail.

Methods

Procedure and participants

Professors working at one of the 8 academic medical centres in The Netherlands were sent an invitational e-mail in September 2011 to participate in a survey addressing burnout symptoms, but also aspects of publication culture. We included professors working in either clinical or preclinical disciplines, all being employed by one of the 8 University Medical Centres in The Netherlands.

The e-mail explained the objectives of the study, using neutral terms as 'work experiences and engagement', and provided them with a link to an anonymous online questionnaire on a protected website. Those who did not respond were sent a reminder after 3 weeks, and responses were registered until 6 weeks after the first invitation.

Variables

The questionnaire contained, apart from demographic questions, validated burnout and engagement questionnaires. Burnout was measured using the Dutch version (25) of the Maslach Burnout Inventory (MBI) Human Services Survey (1), which is designed specifically for use in people working in human services and health care. The Dutch version (the Utrechtse Burn Out Schaal (UBOS), see online appendix for English translation) consists of 20 items covering the three domains of burnout: 1) the depletion of emotional reserves (*emotional exhaustion, 8 items*), 2) an increasingly cynical and negative approach towards others (*depersonalization, 5 items*), and 3) a growing feeling of work-related dissatisfaction (*personal accomplishment, 7 items).*

As examples, emotional exhaustion is assessed through questions such as 'I feel like I am at the end of my rope' and 'I feel burned out by my work', and depersonalization with questions such as 'I feel I treat some of my faculty and residents as if they were impersonal objects'. Personal accomplishment is assessed with questions such as 'I have accomplished many worthwhile things in this job'. Items were rated on a 7-point frequency scale (0-6), such that more points on the emotional exhaustion and depersonalisation domain indicated a higher propensity for having burnout). Personal accomplishment is inversely related to burnout: lower scores indicate a higher propensity for having burnout. Since emotional exhaustion is considered the key component of burnout (1;21-24), we use emotional exhaustion as the primary outcome measure and main variable to assess burnout.

The nominal cut-off scores for burnout were used. These cut-off levels are sometimes based on the Emotional Exhaustion domain scores only. The Dutch Central Bureau of Statistics , for example, has set the cut-off level for the nominal definition of having burnout on an Emotional Exhaustion sub score threshold of >17.68 points (www.cbs.nl, http://www.tno.nl/downloads/Rapport NEA 2010.pdf).

Engagement was measured using the 17-item Utrecht Work Engagement Scale (UWES) (17). This questionnaire has good psychometric properties (18), and consists of three engagement subscales: vigour (6 items), dedication (5 items) and absorption (6 items). High levels of mental energy and willingness to invest in work define vigour, whereas dedication is defined as 'feelings of enthusiasm, pride and inspiration', and absorption implies 'a sense of time passing quickly and low distraction'. Items were rated on a 7-point Likert scale (0-6). The sum of all items is used as a total engagement score.

The demographic and general background information included gender, age (divided into 5 categories), marital status, having homeliving children, type of specialty; years since appointment (per 5 years), main professional activity (research, education, patient care, or management) and self-reported Hirsch Index, a citation-based individual indicator of scientific impact (26).

In this research no patients were involved; therefore no ethics approval was necessary as the research complies with national regulations (https://www.vcmo.nl/wmo/niet-wmo-plichtig-onderzoek/).

Statistical analysis

Analysis of Variance was used to compare groups. Pearson's correlation coefficients were calculated to examine relationships between continuous variables. Multiple linear regression analysis was used to identify independent determinants of burnout scores on a continuous scale. We were cautious to avoid statistical overadjustment with multiple age-related variables. Therefore, we introduced in the first multivariate analysis only demographic and job-specific variables. Variables that conceivably were mediators of effects of demographic and job-specific items were subsequently introduced in a second multiple regression model. In a secondary analysis, logistic regression was performed to analyse the dichotomized burnout scores using cut-off scores. The Statistical Package for the Social Sciences (SPSS) statistics (Chicago USA 2011, version 20) was used for the statistical analyses.

Results

Demographics

Of the 1366 e-mail addresses used, 160 bounced, most often because the addresses no longer existed, or repeatedly provided an out-of-office reply. To the remaining 1206 e-mails; 578 professors responded (49%), of whom 437 (36%) completed the full questionnaire. Data on demographic and job-specific characteristics of complete respondents are summarized in Table 1.

Early-career professors show higher emotional exhaustion scores

Univariate determinants of burnout and engagement (sub)scores on a continuous scale are shown in Table 2. Younger age, less years since appointment, and having children living at home were significantly associated with emotional exhaustion and with at least one other component score of burnout. In the multivariate analysis of demographic and job-specific items, age, home-living children and years since appointment were included, and the latter appeared to be the main, independent age-related determinant of emotional exhaustion (Table 3). We also performed multivariate analysis for depersonalisation in which no effect was found (data not shown).

According to the aforementioned cut-off level on the Emotional Exhaustion scale, 23.8% of medical professors (n = 104) suffered from burnout.

Table 1. Demographic and job-specific characteristics of 437 respondents

Table 1 Demographic and job-specific characteristics of 437 respondents		N = 437
Gender	Male	345 (79%)
	Female	92 (21%)
Age	26-35	1 (0,2%)
	36-45	35 (8%)
	46-55	206 (47%)
	56-65	190 (44%)
	65 and older	5 (1%)
Marital status	Married or cohabiting	401 (92%)
	Single	36 (8%)
Home living children	None	217 (50%)
	1	56 (13%)
	2	96 (22%)
	3 or more	68 (15%)
Years since appointment	0-5	150 (34%)
	6-10	129 (30%)
	11-15	86 (20%)
	16 or more	72 (16%)
Nr. 1 Work priority	Research	255 (59%)
	Education	40 (9%)
	Patient care	63 (14%)
	Management	79 (18%)
Appointment	Temporary	144 (33%)
	Permanent	293 (67%)
Raw scores of burn out dimensions	Emotional Exhaustion Total Score (0-48)	11,9 (SD 8,9)
	Depersonalisation Total score (0-30)	4,4 (SD 4.4)
	Personal accomplishment Total score (0-42)	30,9 (SD 5,9)
Raw scores of engagement dimensions	Vitality Total Score (0-36)	28.1 (SD 5.0)
	Dedication Total Score (0-36)	24.9 (SD 4.2)
	Absorption Total Score (0-36)	26.4 (SD 5,4)
Specialty	Preclinical	81% (354)
	Clinical	16% (70)
	Anonymous	3% 913)

Logistic regression analysis with the nominal burnout outcome variable identified the same determinants as did linear regression analysis for the continuous subscores, albeit with lower levels of statistical significance (data not shown).

The role of the Hirsch index

Among respondents, 74% knew their current Hirsch index (n = 321), and their average index value was 32.6 (standard deviation: 14.9, see Figure 2 for distribution). The Hirsch index was inversely correlated with burnout symptoms, predominantly with the components emotional exhaustion and personal accomplishment, but not with depersonalization. The highest Hirsch index tertile was associated with a significant 19% lower emotional exhaustion score compared to the lower 2 tertiles (Figure 1, panel A). Personal Accomplishment subscore was significantly and more linearly related with the Hirsch index (Figure 1, panel B, beta per tertile 0.1, CI 0.2-1.7). As the H-index is driven by age, these associations were adjusted for age, which did not change the results. To determine whether the Hirsch index (partly) explains the association between being an early career professor and higher burnout scores on the emotional exhaustion domain, multiple regression was performed (Table 4), suggesting that this was not the case (beta per 5-years since professorship changed from -1.3 to -1.5). Also, no statistical interaction between being early career and the Hirsch index was noted (p = 0.8 for the interaction term).

Table 2. Univariate regression analysis comparing independent variables with burnout and engagement component scores (Vigour, Dedication and Absorption)

	Burnout domain scores			Engagement domain scores		
	Emotional Exhaustion	Depersonali-sation	Personal Accomplishment	Vigour	Dedication	Absorption
Age (per 10 years	-1.7**	-0.8**	0.3	0.4	-0.1	-0.6
Gender (female)	0.1	0.5	-0.3	-0.4	-0.3	0.7
Marital status (single)	-1.7	-0.3	1.4	0.7	0.7	0.9
Homeliving children (yes)	2.5**	1.1**	0.6	0.9*	0.5	1.2**
Fixed position (yes)	-1.3	-0.0	1.1*	0.7	0.1	0.2
Years since appointment (per 5 years)	-1.3**	-0.2	0.4*	0.4*	0.2	0.0

Regression coefficients are shown and (borderline) significant values are shown by markation: * $0.05 < p < 0.10$; ** $p < 0.05$. Determinants with a univariate p-value of <0.10 were entered in the multiple regression analyses.

Table 3. Multivariate regression analysis comparing independent variables with emotional exhaustion

	Beta (95% CI)	p-value
Emotional Exhaustion (0-45)		
Age (per 10 years)	-0.3 (-1.9 to 1.3)	0.72
Homeliving children (yes)	1,6 (-0.4 to 3.5)	0.11
Years since appointment (per 5 years)	**-1.0 (-1.9 to -0.1)**	**0.03**

Regression coefficients are shown and significant values are shown in bold ($p < 0.05$)

Figure 1. The distribution of the H-index.

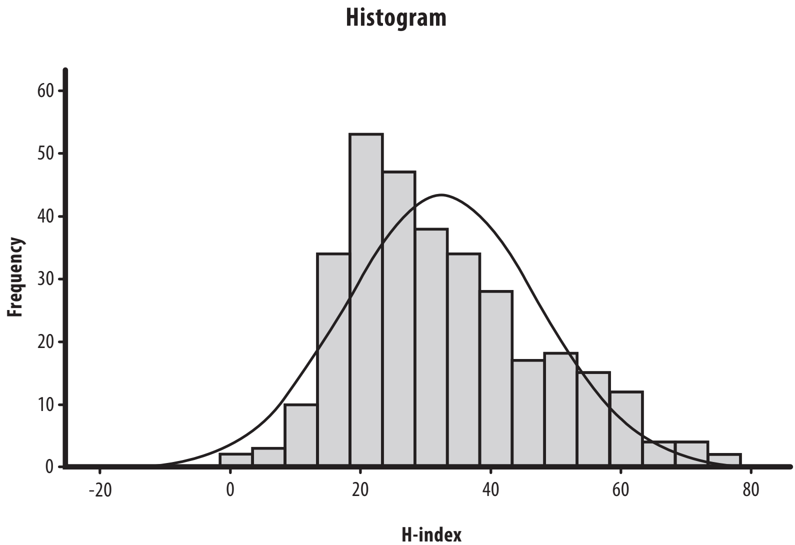
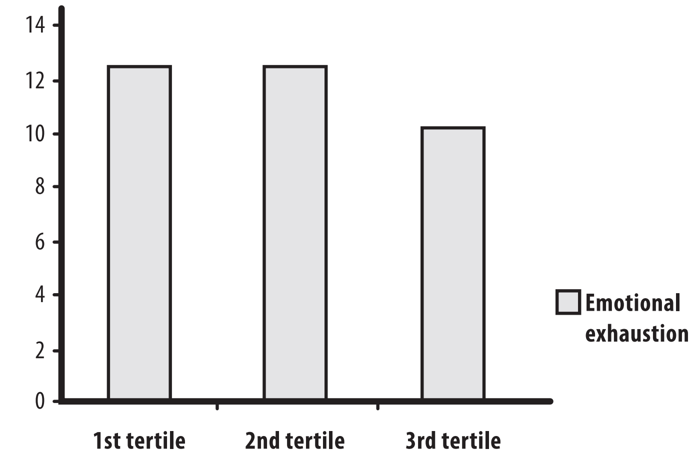


Table 4. Crude and multivariate analysis of emotional exhaustion including the Hirsh-index as additional independent variable.

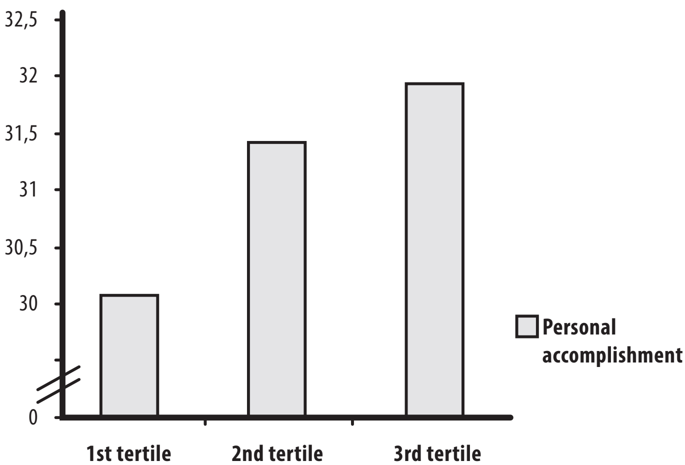
	Beta (CI 95%)	p-value
Emotional Exhaustion (0-45)		
Crude analysis		
Years since appointment (per 5 years)	-1.3 (-2.1 to -0.6)	<0.01
Hirsch index (upper vs lower 2 tertiles)	-2.3 (-4.4 to -0.3)	0.02
Multivariate analysis		
Years since appointment (per 5 years)	-1.5 (-2.5 to -0.6)	<0.01
Hirsch index (upper vs lower 2 tertiles)	-1.2 (-3.3 to 1.0)	0.28

Figure 2. Score on emotional exhaustion (panel A) and personal accomplishment (panel B), divided in tertiles in h-index score. 1st tertile h-index ranging from 10-23, 2nd tertile from 24-37, third tertile ranging from 38-78).

Panel A



Panel B



Burn out and engagement

Vigour and dedication were negatively associated with emotional exhaustion (correlation coefficient -0,36 and -0,38, respectively, both p < 0,001), and to depersonalisation (-0,27 and -0,35, respectively, both p < 0,001). All three subscales of engagement (vigour, dedication and absorption) were positively and strongly related to personal accomplishment (0,61, 0,56 and 0,45, respectively, all p < 0,001).

Furthermore, all three engagement subscales showed significant relations with the Hirsch-index (in tertiles, beta's (CI) 1.1 (0.5 to 1.7), 0.8 (0.3 to 1.2), and 0.9 (0.2 to 1.5), respectively, all p < 0.01).

Discussion

This study suggests that emotional exhaustion is frequent among medical professors, and that the early career years represent a risk period for emotional exhaustion. Having reached a certain degree of scientific success, as indicated by a high Hirsch factor, may confer some degree of protection.

Interpretation of results

In comparable studies, high burnout frequencies were found in academic chairs in specific medical fields such as gynaecology and orthopaedic surgery (3;6). In these studies, 75% of orthopaedic surgeons had moderate to high levels of emotional exhaustion and 54% of gynaecologists reported high levels of burnout (these studies were using different cut-off values compared to this study). We found no previous study addressing an entire nationwide population of medical professors.

Higher emotional exhaustion subscores are found among younger professors, who usually are at the start of their career, and more often have children living at home (not after multivariate analysis). This is in line with previous studies, which found high emotional exhaustion in younger chairs and those with a spouse and children.

Burnout was also more common in new professors (12-15). These three determinants have a high degree of co-linearity, and may be in each other's causal pathway. Therefore, the multivariate analysis, which demonstrated that the number of years since appointment is the prime, independent determinant of burnout symptomatology, should be interpreted with some caution. Feeling of control over work and spouse support are two important protective factors against burnout. Effects of seniority may be explained via these effects, since professional experience may increase the (sense of) control over work and work hours (27). A possible survivor bias is conceivable but not very likely since professors leaving their position in their early years are very rare.

We analysed the potential correlation of the Hirsch index with burnout symptoms separate from demographic and job-specific characteristics. A higher Hirsch index was related to lower emotional exhaustion scores, but did not explain, at least not statistically, the impact of being early in a professor career on burnout. Whether a low Hirsch index causes extra stress, or a high Hirsch index is a protective factor is a semantic, or even philosophical issue. In terms of career chances in academia, the Hirsch index may be a stressor for youngsters, but could also be reassuring for seniors. Furthermore the Hirsch index is correlated to personal accomplishment and all three subscales of engagement. Apparently, medical professors with a high Hirsch index feel they are more capable, have more vigour and dedication, and are more absorbed in their work.

Finally, engagement correlates moderately with burnout subscale scores. The interpretation of these correlations is hampered by the likelihood that the causality is bidirectional: engagement may protect against burnout, and burnout can severely compromise engagement. There may exist independent effects, but only longitudinal research can address this.

Strengths and limitations

The strength of this study is that the survey was nationwide, addressing all medical professors in the country. Furthermore, burnout domain subscores were analysed on their natural, continuous scale, avoiding the loss of power associated with (arbitrary) dichotomisation of burnout symptomatology. In this respect, the topic of our study is more the propensity for developing burnout, rather than qualifying for any formal definition of the disorder.

A number of limitations also need to be addressed. First, we cannot rule out response bias. The survey completion rate was 36%, which is comparable to similar types of online questionnaires (28). Although response bias is difficult to investigate, it is interesting to speculate in which direction it would occur. We think that response bias in our study may in fact have been bidirectional. Those experiencing more burnout symptoms could either preferentially participate (identification with the topic) or be reluctant to do so, caused by a sense of lack of time and task overload. To assure the representativeness of the sample we investigated the distribution of age and gender among all professors in the Netherlands. This population was representative as +/- 17% of the medical professors in the Netherlands is female (our sample 21%, http://www.stichtingdebeauvoir.nl/wp-content/uploads/MonitorVrouwelijke_Hoogleraren2012.pdf) and the average age of professors in the Netherlands in another study including 1256 professors, was comparable with our mean ages (29). This supports representativeness of our study population.

Since all medical professors in the Netherlands were invited for participation, the responders are not a sample from a sample but a sample of total study population. This further supports the representativeness of the study population. The population of medical professors is, obviously, heterogeneous. In the Netherlands, most have, at least formally, a part-time appointment as professor. All are more intensively involved in management, research and educational activities than regular physicians, but the degree to which this is the case may vary. Importantly, all professors in the Netherlands spend at least 1 to 2 days on patient care in view of registration legislation.

There may also be a taboo on burnout, causing respondents to downplay the severity and personal impact of burnout, despite the fact that anonymity was guaranteed. Another potential limitation could be the use of an online questionnaire for such a sensitive issue. However, the validity of online questionnaires is probably similar to 'live' questionnaires. (30) The timing of the study (September-October) could also have influenced the results and possibly attenuate burnout symptom scores, since national holidays are held in July and August, and academic work normally starts in the beginning of September.

A final important issue is the risk of framing: creating an atmosphere which stimulates 'positive answers' depending on how the topic is introduced, how the questions are phrased, etc. To limit this risk, the invitation e-mail did not contain words such as 'burnout', but was phrased using more neutral words as 'work engagement' and 'job satisfaction' The Maslach Burnout Inventory is also constructed to reduce this risk of framing by including positive questions in the domain of personal accomplishment, which improves psychometric properties (31).

Maslach's definition of burnout was originally a division of a sample into equal thirds and cut-off values were not mentioned. Burnout as a domain is most often defined as being above cut off on at least two dimensions (high emotional exhaustion and depersonalisation or high emotional exhaustion and low personal accomplishment). Since the Dutch bureau of statistics provides cut-off values for emotional exhaustion only, emotional exhaustion was, with possible limitation, chosen as a core feature of burnout. Furthermore since our study population consisted solely of ambitious and highly skilled medical professors the degree of burnout on the personal accomplishment domain was extremely low and was therefore considered not to be an appropriate feature of measuring burnout in this population.

Therefore we chose not to compare with other thresholds since in different research different thresholds are used and are therefore ambiguous and inconclusive (32;33).

We also used cutoff values of the Dutch national Central Bureau of Statistics to allow a comparison with other Dutch professionals. In national samples of the total working force in the Netherlands, 11-14% meet the criteria for burnout. In another sample among Dutch doctors in residency training programs, the percentage was 41%, using the same definition as we did. However, the CBS assesses burnout using 5 statements from the emotional exhaustion scale to define moderate or severe burnout. Hence, these comparisons suggests that being a resident is be more stressful than being a professor, and that both are more prone to burn-out than the general Dutch working population.

Finally, the fact that we did not include other potential burnout determinants such as weekly work hours or work-home conflicts precludes more detailed analyses of the wider spectrum of determinants of burnout among this group.

Conclusion

We conclude that emotional exhaustion is common among Dutch medical professors, and are determined by several factors, all related to being in an early stage of their professional career. Further research should focus on the impact of burnout on both the personal level, as well as on the level of professional performance in the clinical, educational and scientific domains. In future studies, potential preventive strategies should be addressed.

Reference List

1. Maslach C, Jackson SE, Leiter MP: *Maslach Burnout Inventory Manual.* 3rd edition. Palo Alto, CA: Consulting Psychologists Press; 1996.
2. Cruz OA, Pole CJ, Thomas SM: **Burnout in chairs of academic departments of ophthalmology.** *Ophthalmology* 2007;**114:**2350–2355.
3. Gabbe SG, Melville J, Mandel L, Walker E: **Burnout in chairs of obstetrics and gynecology: diagnosis, treatment, and prevention.** *Am J Obstet Gynecol* 2002;**186:**601–612.
4. Johns MM III, Ossoff RH: **Burnout in academic chairs of otolaryngology: head and neck surgery.** *Laryngoscope* 2005;**115:**2056–2061.
5. Prins JT, Hoekstra-Weebers JE, Gazendam-Donofrio SM, Dillingh GS, Bakker AB, Huisman M, et al. **Burnout and engagement among resident doctors in the Netherlands: a national study.** *Med Educ* 2010;**44:**236–247.
6. Saleh KJ, Quick JC, Conaway M, Sime WE, Martin W, Hurwitz S, *Einhorn TA*: **The prevalence and severity of burnout among academic orthopaedic departmental leaders.** *J Bone Joint Surg Am* 2007;**89:**896–903.
7. Shanafelt TD, Boone S, Tan L, Dyrbye LN, Sotile W, Satele D, West CP, Sloan J, Oreskovich, MR: **Burnout and satisfaction with work-life balance among US physicians relative to the general US population.** *Arch Intern Med* 2012;**20:**1–9.
8. Prins JT, van der Heijden FM, Hoekstra-Weebers JE, Bakker AB, van de Wiel HB, Jacobs B, *et al.* **Burnout, engagement and resident physicians' self-reported errors.** *Psychol Health Med* 2009;**14:**654–666.
9. Spickard A Jr, Gabbe SG, Christensen JF: **Mid-career burnout in generalist and specialist physicians.** *JAMA* 2002;**288:**1447–1450.
10. West CP, Huschka MM, Novotny PJ, Sloan JA, Kolars JC, Habermann TM, et al. **Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study.** *JAMA* 2006;**296:**1071–1078.
11. West CP, Tan AD, Habermann TM, Sloan JA, Shanafelt TD: **Association of resident fatigue and distress with perceived medical errors.** *JAMA* 2009;**302:**1294–1300.
12. Demerouti E, Bakker AB, de JJ, Janssen PP, Schaufeli WB: **Burnout and engagement at work as a function of demands and control.** *Scand J Work Environ Health* 2001;**27:**279–286.
13. Martini S, Arfken CL, Churchill A, Balon R: **Burnout comparison among residents in different medical specialties.** *Acad Psychiatry* 2004;**28:**240–242.
14. Maslach C, Jackson SE: **The role of sex and family variables in burnout.** *Sex Roles* 1985;**12:**837–851.
15. Maslach C, Schaufeli WB, Leiter MP: **Job burnout.** *Annu Rev Psychol* 2001;**52:**397–422.

16. Rovik JO, Tyssen R, Hem E, Gude T, Ekeberg O, Moum T, et al. **Job stress in young physicians with an emphasis on the work-home interface: a nine-year, nationwide and longitudinal study of its course and predictors.** *Ind Health* 2007;**45:**662–671.
17. Schaufeli WB, Bakker AB: *UBES: Utrechtse Bevlogenheidschaal [UBES: Utrecht Work Engagement Scale].* Utrecht: University of Utrecht; 2003.
18. Schaufeli, Salanova M, Gonzalez-Roma V, Bakker AB: **The measurement of engagement and burnout: a two-sample confirmatory factor analytic approach.** *J Happiness Stud* 2002;**3:**71–92. .
19. Dyrbye LN, Shanafelt TD, Balch CM, Satele D, Sloan J, Freischlag J: **Relationship between work-home conflicts and burnout among American surgeons: a comparison by sex.** *Arch Surg* 2011;**146:**211–217.
20. Shanafelt TD, Balch CM, Bechamps GJ, Russell T, Dyrbye L, Satele D, et al. **Burnout and career satisfaction among American surgeons.** *Ann Surg* 2009;**250:**463–471.
21. Bekker MMJ, Croon MA, Bressers B: **Childcare involvement, job characteristics, gender and work attitudes as predictors of emotional exhaustion and sickness absence.** *Work & Stress* 2005;**19:**221–237.
22. Brenninkmeijer V, VanYperen N: **How to conduct research on burnout: advantages and disadvantages of a unidimensional approach in burnout research.** *Occup Environ Med* 2003;**60:**i16–i20.
23. Schaufeli WB, Taris TW: **The conceptualization and measurement of burnout: Common ground and worlds apart.** *Work Stress* 2005;**19:**256–262.
24. Shirom A: **Job related burnout**. In *Handbook of Occupation Health Psychology.* Edited by J.C.Quick &LET. Washington DC: American Psychological Association; 2003:245–265.
25. Schaufeli WB, van Dierendonk D: *Utrechtse Burnout Schaal (UBOS), Handleiding [Utrecht Burnout Scale, Manual].* Utrecht: Swets & Zeitlinger; 2000.
26. Hirsch JE: **An index to quantify an individual's scientific research output.** *Proc Natl Acad Sci USA* 2005;**102:**16569–16572.
27. Keeton K, Fenner DE, Johnson TR, Hayward RA: **Predictors of physician career satisfaction, work-life balance, and burnout.** *Obstet Gynecol* 2007;**109:**949–955.
28. Cook C, Heath F, Thompson RL: **A meta-analysis of response rates in web- or internet-based surveys.** *Educ Psychol Meas* 2000;**60:**821–836.
29. Gerritsen M, Sanden vJ, Verdonk T, Visser A: *Monitor vrouwelijke hoogleraren*; 2012.
30. Joinson A: **Social desirability, anonymity and Internet-based questionnaires.** *Behav Res Methods Instrum Comput* 1999;**31:**433–438.
31. Streiner DL, Norman GR: *Health Measurement Scales : A Practical Guide to Their Development and use.* Oxford; Oxford University Press; 2008.

32. Orena EF, Caldiroli D, Cortellazzi P: **Does the Maslach Burnout Inventory correlate with cognitive performance in anesthesia practitioners? A pilot study.** *Saudi J Anaesth* 2013;**7:**277–282.

33. Prosser D, Johnson S, Kuipers E, Szmukler G, Bebbington P, Thornicroft G: **Mental health, "burnout" and job satisfaction among hospital and community-based mental health staff.** *Br J Psychiatry* 1996;**169:**334–337.

4

Chapter 4

Publication pressure and burn out among Dutch medical professors: a nationwide survey

4

J.K. Tijdink, A.C.M. Vergouwen, Y.M. Smulders

PLoS One 2013;8:e73381. doi:10.1371/journal.pone.0073381

Abstract

Background:

Publication of scientific research papers is important for professionals working in academic medical centres. Quantitative measures of scientific output determine status and prestige, and serve to rank universities as well as individuals. The pressure to generate maximum scientific output is high, and quantitative aspects may tend to dominate over qualitative ones. How this pressure influences professionals' perception of science and their personal well-being is unknown.

Methods and Findings:

We performed an online survey inviting all medical professors (n = 1206) of the 8 academic medical centres in The Netherlands to participate. They were asked to fill out 2 questionnaires; a validated Publication Pressure Questionnaire and the Maslach Burnout Inventory. In total, 437 professors completed the questionnaires. among them, 54% judge that publication pressure 'has become excessive', 39% believe that publication pressure 'affects the credibility of medical research' and 26% judge that publication pressure has a 'sickening effect on medical science'. The burn out questionnaire indicates that 24% of medical professors have signs of burn out. The number of years of professorship was significantly related with experiencing less publication pressure. Significant and strong associations between burn out symptoms and the level of perceived publication pressure were found. The main limitation is the possibility of response bias.

Conclusion:

A substantial proportion of medical professors believe that publication pressure has become excessive, and have a cynical view on the validity of medical science. These perceptions are statistically correlated to burn out symptoms. Further research should address the effects of publication pressure in more detail and identify alternative ways to stimulate the quality of medical science.

Background

Publication of scientific research papers is important for medical professionals, particularly in academic environments. Scientific output is used to rank prestige and status of both academic medical centres as well as of individual medical staff (1). Quantitative measures of scientific performance, such as the Hirsh index (2), have become particularly important, as these directly influence grant proposals, financial rewards and career potential (3–8).

Parallel to increased emphasis on scientific output measures, the quantity of (medical) scientific output has increased enormously in the past decades: the number of scientific journals increased from 5000 in 1997 tot 8000 journals in 2010 as registered by the ISI Web of Knowledge. (ISI Web of Knowledge, http://www.webofknowledge.com, consulted March 2012), and the amount of scientific papers doubles every 12 years (http://www.scopus.com).

The increasing emphasis of scientific performance potentially raises pressure on medical professionals to publish, and may intensify competition between them. This competition for papers and funding is often considered a salutary driving force among scientists, increasing efficiency and productivity (9). Potential negative effects of a competitive publication culture with a focus on quantitative performance are not often considered. Concerns have however been expressed that scientists are continuously producing 'publishable' results at the expense of quality, validity, scientific rigour and personal integrity (10), and published negative research results have decreased over the years (11). Consequently, clinical practice based on research outcomes may be jeopardised (12;13) .

Excessive emphasis on scientific output may also affect academic activities that compete with science for time and attention, such as clinical and educational activities (14), as these activities can be perceived disadvantageous and less important while affected by publication pressure (13;15).

Finally, it is conceivable that mental well-being benefits from working in an environment with a healthy scientific culture. In this respect, increasing evidence that burn out symptoms are common among medical doctors and residents (16–20) is noteworthy. Burn-out symptoms may impact on academic tasks (ie not just science, but also patient care and education) and has previously been suggested to be related to publication pressure (13).

Many of the aforementioned phenomena are difficult to address in quantitative terms and epidemiological studies. However, important information can be obtained from anonymous questionnaires in academic professionals. The aim of our study is therefore to assess the perception of publication pressure among a large group of medical professionals in an advanced stage of their academic career. We also investigated how publication pressure relates to their view on the quality of medical science, as well as to aspects of their personal well-being.

Methods

Procedure and participants

Full professors working at one of the 8 academic medical centres (AMC's) in The Netherlands were sent an invitational e-mail in September 2011.

The deans of 4 AMC's provided the e-mail addresses of their professors (n=600) to the research team. The other 4 AMC's chose to distribute the electronic link to the questionnaire internally.

The invitational e-mail explained the objectives of the study, using neutral terms as 'work engagement' and 'publication culture', and provided them with a link to an anonymous online questionnaire on a protected website. Those who not responded were sent a reminder after 3 weeks.

Variables

The questionnaire contained, apart from demographic data, 2 parts; a specifically designed publication pressure questionnaire and a validated burn out questionnaire.

The Publication Pressure Questionnaire (PPQ) contained 24 statements (table 2), the responses to which were scored on a 5-point Likert scale. The questionnaire

Table 1. Publication pressure questionnaire (PPQ). In **bold** the statements who are part of the validated questionnaire. * inversed questions; higher scores for disagreement

Questions (Domain)	Likert Scale Score (SD)	% agreement (4-5 or 1-2* on Likert Scale)
1. Without publication pressure, my scientific output would be of higher quality (PP1)	2,6 (1,2)	23 %
2. Publishing scientific articles is the most important part of my professorship (PP1)	3,2 (1,1)	41 %
3. My scientific publications contribute to better (future) medical care* (PP1)	2,2 (0,8)	74 %
4. The number of scientific publications determines my status/prestige (PP3)	4,0 (0,8)	79 %
5. The number of scientific publications of my colleagues determines their status/prestige (PP2, PP3)	4,0 (0,8)	82 %
6. I experience judgement of my publications by colleagues as stressful (PP1, PP3)	2,4 (1,1)	18 %
7. I experience the scientific output criteria set by the university for my appointment and reappointment as stimulating* (PP1, PP3)	3,2 (1,1)	24 %
8. Publication pressure puts pressure on my relations with fellow researchers (PP1,PP2)	2,8 (1,2)	32 %
9. Publication pressure increases my scientific output, without loss of quality* (PP1)	3,5 (1,0)	15 %
10. I suspect that in some colleagues publication pressure leads to (un)intentional data manipulation (PP2)	3,0 (1,1)	33 %
11. Worldwide, publication pressure adds validity to medical science* (PP2)	3,7 (1,0)	11 %
12. On a global scale, publication pressure causes serious doubts regarding the validity of research results (PP2)	3,1 (1,1)	38 %
13. I think the pressure to publish has become excessive (PP1, PP2))	3,4 (1,2)	54 %
14. Fellow medical experts envy me because of my professorship (PP3)	3,0 (0,9)	13 %
15. The competitive scientific culture stimulates me to publish more* (PP1)	3,0 (1,0)	40 %
16. I experience my professorship as a burden and I sometimes long back to when I was not in this position (PP3)	2,0 (1,1)	13 %
17. My colleagues mainly judge me on my publication record (PP1, PP3)	2,6 (1,1)	22 %
18. Despite the pressure to publish, I enjoy investing in other activities which come with my professorship* (PP1)	1,9 (0,8)	85 %
19. Fellow professors adequately maintain their clinical and educational skills, despite publication pressure* (PP2)	2,9 (0,9)	36 %
20. Team spirit and collegiality are, in my hospital, key aspects in all professors' appointment procedures* (PP3)	3,0 (1,1)	34 %
21. I cannot trust my colleagues on innovative research proposals (PP1, PP3)	2,1 (1,0)	11 %
22. I am too much involved in management (PP3)	3,2 (1,2)	43 %
23. Professorship is difficult to combine with training and educating residents (PP3)	2,9 (1,0)	25 %
24. The urge to publish makes science sick (PP2)	2,6 (1,2)	25 %

was initially divided into 3 broad domains by intuition:

1) pressure to publish personally experienced by the respondent
2) publication pressure in general terms in the academic work place, as perceived by the respondent
3) publication pressure relating to the scientist's position and status (e.g. promotion, re-appointment, etc)

After statistical validation of the PPQ with Confirmatory Factor Analysis, Explanatory Factor Analysis and Item Response Theory, we condensed the questionnaire to 14 items, having publication pressure as a single factor (21).

The Likert scale scores were assigned 1 to 5 points such that higher scores reflected higher pressure. We labelled different statements as positive and negative to avoid 'yeah-saying' (22). Negative statements were scored inversely (ie totally disagree = 5 points instead of 1 point, see table 1)

Burn out was measured using the Dutch version (23) of the Maslach Burnout Inventory (MBI) Human Services Survey (24) , which is designed specifically for use in people working in human services and health care. We chose the MBI since it is the primary measurement for work-related mental status in otherwise healthy people. Also, the MBI is most frequently used in similar types of scientific research (16–20).

The Dutch version of the MBI consists of 20 items covering the three domains of burn out: emotional exhaustion (EE, 8 items, key symptom), depersonalisation (DP, 5 items) and personal accomplishment (PA, 7 items). Emotional exhaustion (EE) is characterized by loss of energy at work and a negative attitude towards work-related activities. Depersonalisation (DP) relates to a kind of alienation from work, where interest in job and colleagues is lost.

Personal accomplishment (PA) is a positive symptom and reflects feelings of capability and job satisfaction.

Items were rated on a 7-point frequency scale and assigned 0-6 points, such that more points indicated a higher propensity for having burn out. Cut-off scores for 'having a burn out' were provided by the Dutch Central Bureau of Statistics (www.cbs.nl).

Table 2. Demographics

Demographics			
		N=437	**%**
Gender	Male	345	79
	Female	92	21
Age	26-35	1	0,2
	36-45	35	8,0
	46-55	20	47,1
	56-65	190	43,5
	65 and older	5	1,1
Marital status	Married or cohabiting	401	92
	Single	36	8
Home living children	None	217	50
	1	56	13
	2	96	22
	3 or more	68	15
Years of professorship	0-5	150	34
	6-10	129	30
	11-15	86	20
	15 or more	72	16
Nr. 1 Work priority	Research	255	59
	Education	40	9
	Patient care	63	14
	Management	79	18
Appointment	Temporary	144	33
	Permanent	293	67

Respondents provided demographic information on gender, age, marital status, children living at home; type of specialty; years of professorship, and main professional activity (research, education, patient care or management).

Statistical Analysis

Independence of bivariate association was assessed by multiple linear regression analysis using the stepwise forward method. Pearson's correlation coefficients were calculated to examine relationships between publication pressure and burn out scores.

An age-adjusted General Additive Model (GAM)-curve was constructed to graphically display the association between publication pressure and specific burn out scores (25), using the statistical software package R, version 2.15.1 (R Development Core Team, R foundation, USA).

Results

In total, we used 1366 e-mail addresses. Of these, 160 bounced, most often because the addresses no longer existed, or provided an out-of-office reply. Of the 1206 professors left; 578 responded (49%) , of whom 437 (36%) completed the full questionnaire. The demographic data of the complete responders are summarized in table 2.

Publication Pressure Questionnaire (PPQ)

Table 1 lists the questions and responses. The responses to some of the key questions indicate that the majority (54%, item 13) rates publication pressure as 'excessive'. In addition, 1 out of every 3 to 4 respondents (items 1,10 and 12) believes that the pressure to publish has detrimental effects on the validity and credibility of medical science. Finally, 24% (item 24) qualifies publication pressure as having a 'sickening' effect on medical science.

We use the sum score of the 14 items (marked as **bold** in table 1) of the condensed, validated PPQ for further analysis of correlations and determinants.

Table 3. Univariate and Multivariate analysis comparing independent variables with the PPQ.

Univariate	Beta	CI 95% lower bound	CI 95% upper bound	P value
Age (10 year)	-0.623	-1.75	0.51	0.28
Gender (female)	1.851	0.04	3.89	0.05
Marital status (single)	-1.421	-4.12	1.78	0.30
Homeliving children (increase of 1 child)	0.718	0.08	1.36	0.03
Fixed position (yes)	-1.116	-2.69	0.46	0.17
Years prof (increase of 5 years)	-1.013	-1.69	-0.33	0.004
Multivariate				
Gender (female)	1.528	-0.30	3.36	0.10
Homeliving children	0.479	0.17	-0.20	0.17
Years of professorship	-0.739	-1.48	-0.003	0.05

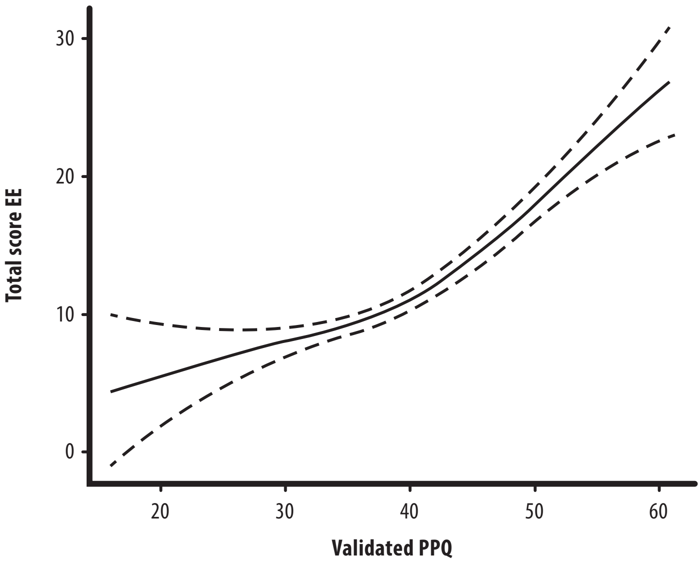


Figure 1. General Additive Model-curve, demonstrating the association between the sum score on the validated 14-item publication pressure questionnaire, and the Emotional Exhaustion component of the burn out index, adjusted for gender to reduce residual confounding (Beneditti A and abrahamowicz M). The dotted line represents the 95% Confidence interval.

Correlates of publication pressure

In univariate analyses, female gender and having home living children were positively, and the number of years of professorship was negatively associated with publication pressure (table 3, top).

Multivariate analysis identified the number of years of professorship as the single independent (inverse) determinant of publication pressure. The univariate effects of gender and having home living children were reduced to non-significant trends (table 3, bottom).

Prevalence of burn out and Association with Publication Pressure Questionnaire

Of all respondents, 24 % met the formal Dutch CBS criteria for having burn out. Mean scores of EE, DP and PA were 11.9 (interquartile range (IQR) 5-16), 4.4 (IQR 1-6) and 30.9 (IQR 28-35).

The PPQ-score was significantly associated with scores on all 3 subscales of the burnout questionnaire. Correlations were strongest for the emotional exhaustion subscale (Pearson's correlation coefficient 0.45, $p<0.001$). The cumulative publication pressure score showed a weaker, but highly significant (inverse) correlation with depersonalisation and personal accomplishment (Pearson's correlation coefficient 0.29 and -0.15 respectively, $p<0.001$). Figure 1 illustrates the gender-adjusted co rrelation between the PPQ-score and the emotional exhaustion subscale of the burnout questionnaire.

Discussion

Our study suggests that a substantial proportion of the responding medical professors judge publication pressure as having become excessive, and a substantial part believes that this affects the validity and credibility of medical science. Furthermore, they personally experience publication pressure, and believe the publication pressure negatively influences their work both in science as well as in other academic tasks, such as clinical and educational work. There is a high level of burn out among medical professors in The Netherlands, and publication pressure correlates positively with burn out symptoms, particularly in the domain of emotional exhaustion.

Our findings are generally in line with the few studies published on this subject, which relate publication pressure to publication bias (26), and report high personal pressure associated with academic competition (12) . Our results add to these previous studies in that our survey is nationwide, identify possible determinants of experienced publication pressure, and address associations of such pressure with perception and trust in medical science. Finally, we addressed personal consequences of publication pressure in terms of burn-out symptoms.
The causes of publication pressure are important to consider. Although our study was not specifically designed to identify such causes, the responses to statements #4, 5 and 15 in the questionnaire suggest that increasing emphasis of quantitative aspects of scientific output plays an important role. This contention is shared by others, who have published on the effects of introduction of the Hirsch-index on the medical scientific field (12); (27;28).

Another important cause may be the importance of bibliographic parameters in the assignment of research funding, grants, scholarships and academic positions. In this area, ambition and prestige is partly built on bibliographic parameters, potentially compromising the impact of clinical and educational performance.

Publication pressure may have adverse effects on medical science. In the respondents' opinion, publication pressure can adversely affect validity and reliability of the medical literature. A bias towards positive outcomes has indeed been suggested in increasingly competitive, academic environments (26). Competitiveness and precariousness of scientific careers have increased, and evidence that this might contribute to scientific bias and even misconduct has accumulated (29). Scientists in focus groups suggested that the need to compete in academia is a threat to scientific integrity (12). Those found guilty of scientific misconduct often invoke excessive pressure as part of the explanation for their actions (12;27;30;31). Surveys suggest that competitive research environments decrease the likelihood to follow scientific ideals and increase the likelihood to witness scientific misconduct (27;32) Another potential adverse effect is neglect of other important academic skills, such as education and patient care.

Apart from general effects on medical science and practice, our study suggests that excessive publication pressure has detrimental effects on personal well-being. A large proportion of the respondents experiences publication pressure as a burden. Moreover, our study indicates that publication pressure causes them to develop a cynical view on medical science, and may be associated with increased risk of developing burn-out.

The central Bureau of Statistics Netherlands has reported that 8-11% of the Dutch working population is burned out (www.cbs.nl). This suggests that burn out is more common among medical professors (24%), which is in line with previous reports (16–18;20). Burn out among medical professionals not only causes personal suffering, but also leads to decreased work performance and jeopardises the quality of patient care (33;34).

A number of limitations of our study needs to be addressed. Firstly, with a response rate of 36%, we cannot rule out response bias. Nevertheless, this response rate of 36% is normal for internet-based surveys in the general population (35) and for similar surveys among academics (36) http://www.supersurvey.com/papers/supersurvey_white_paper_response_rates.pdf).

Although the response rate of 36% could be considered average and ideally would have been higher, it does represent a sample of the total population of medical professors in the Netherlands, were we reached all medical professors, not of a sample of them.

Regardless of whether a higher response rate would have been theoretically feasible, it is prudent to discuss the direction of potential sources of response bias. We think that response bias in our study may have been bidirectional; non-response may be related to lack of time or sense of task overload. This is demonstrated by Prins and colleagues (19) who sent out an ultra brief questionnaire to all the nonresponding residents asking for reasons of lack of cooperation. 23% of non-respondents did not participate because of lack of time and 11 % did not because of lack of energy. Similar reasons for non-response in our survey may have caused underestimation of burn-out symptoms and of discontent with publication pressure. Such underestimation might also have resulted from a taboo on personal pressure and burn out, causing respondents to downplay the severity and personal impact of publication pressure, despite guaranteed anonymity for the respondents.

On the other hand, lack of cooperation may be related to the subject of the survey. Possibly, some medical professors consider publication pressure irrelevant, and thus refuse to participate. Such bias would conceivably have caused overestimation of burn-out and disapproval of publication pressure among respondents. Another potential source of overestimation of the problem is framing. The invitation e-mail did not contain words as 'burn out' and 'publication pressure' but was phrased using more neutral words as 'work engagement' and 'publication culture', to increase response rate and avoid framing and related response bias as much as possible. Some of the statements we used for the questionnaire were perhaps 'framed', but we believe questionnaire statements should connect to the context of the ongoing academic debate on the subject of publication pressure, which is framed in itself. We therefore chose to include provocative 'negative' statements/questions, but the dominance of such negative statements persuaded us to mix them with more positively framed statements/questions. The 'negative' questions/statements connect to the current public and academic debate, and are expected to be recognised as such. The 'positive' ones are those that provoke respondents to take the opposite, or at least a more reflective, position. Inclusion of such inversed 'positive' statements improves psychometric properties of the questionnaire and downplays yeah-saying (22) (see methods).

Burn out was measured using the Maslach Burnout Inventory (MBI). This inventory has its advantages and shortcomings. The MBI is regularly used in similar types of studies addressing work-related psychological stress in medical professionals. Burn-out is not a specific, but in our view a conceivable consequence of publication pressure, which is indeed supported by the literature (13). For further research, the use of alternative job-stress questionnaires or surveys for measuring psychiatric symptoms should be considered.

Another potential source of bias, in either direction, could be the use of an online questionnaire. The validity of online questionnaires has, however, been extensively studied and there is no evidence that web based questionnaires are less valid than 'live' questionnaires (37).

Finally, the timing of the study (September-October 2011) could have influenced the results and possibly attenuate burn out symptom scores, since national holidays are held in July and August, and academic work normally starts in the beginning of September.

We chose to include full medical professors who are in various ways the leading authorities in their research field and have the key positions in the hierarchy in Dutch academia in practising, steering and teaching science. They are role models for younger scientists in the Dutch medical Universities. Therefore, their opinion on publication culture is of main importance. Also, this particular group of medical professionals was not previously examined in a similar way. We can only speculate how other physicians or researchers in academic medical centres, would have responded to the questionnaires. Pressure could be worse among individuals from lower hierarchical groups, who need scientific output to boost their careers, which may predispose them to experience pressure and stress related symptoms (19). Furthermore, our data suggests that accumulation of years of professorship is an important determinant for lower publication pressure scores. Is this light, it is conceivable that being in the early years of academic career development as such is associated with higher levels of publication pressure.

In conclusion, there is a high level of experienced publication pressure in medical professors. No less than 24% of them meet the criteria for having a burn-out, which is statistically correlated to reported publication pressure. These results shed a dim light on some inner thoughts and feelings of our academics leaders in science, education and patient care. Further research is obviously needed, but actions to address the upcoming 'more is better' culture in medical science already appear necessary.

Reference List

1. Garfield E. Citation indexes for science; a new dimension in documentation through association of ideas. Science 1955;122:108-111.
2. Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci 2005;102: 16569-16572.
3. Bedeian AG, Van Fleet DD, Hyman HH. Scientific achievement and editorial-board membership. Organizational Research Methods 2009;12: 238.
4. Bird SJ. Research ethics, research integrity and the responsible conduct of research. Science and Engineering Ethics 2006;12: 412.
5. Gannon F. The impact of the impact factor. EMBO Rep 2000;1: 293.
6. Glick WH, Miller CC, Cardinal LB. Making a life in the field of organization science. Journal of Organizational Behavior 2007;28: 817-835.
7. Retzer V, Juranski G. Towards objectivity in research evaluation using bibliometric indicators - A protocol for incorporating complexity. Basic and Applied Ecology 2009;10: 393-400.
8. Hessels L.K. Science and the struggle for relevance [dissertation]. 2010; Utrecht University.
9. Feller I. The determinants of research competitiveness among universities. In: Teich AH, editors. Competitiveness in academic research. 1996 Washington: American Association for the advancement of Science.
10. Adler N, Harzing AW. When Knowledge Wins: Transcending the Sense and Nonsense of Academic Rankings. Academy of Management Learning and Education 2009;8: 72-95.
11. Fanelli D. Negative results are disappearing from most disciplines and countries. Scientometrics 2012;90: 891-904.
12. Anderson MS, Ronning EA, De Vries R, Martinson BC. The perverse effects of competition on scientists' work and relationships. Sci Eng Ethics 2007;13: 437-461.
13. Miller AN, Taylor SG, Bedeian AG. Publish or perish: academic life as management faculty live it. Career development international 2011;16: 422-445.
14. DeAngelis CD. Professors not professing. JAMA 2004;292: 1060-1061.
15. van Dalen HP, Henkens K. Intended and unintended consequences of a publish-or-perish culture: A worldwide survey. Journal of the American Society for Information Science and Technology 2012; 63: 1282-1293.
16. Cruz OA, Pole CJ, Thomas SM. Burnout in chairs of academic departments of ophthalmology. Ophthalmology 2007; 114: 2350-2355.
17. Gabbe SG, Melville J, Mandel L, Walker E. Burnout in chairs of obstetrics and gynaecology: diagnosis, treatment, and prevention. Am J Obstet Gynecol 2002;186: 601-612.

18. Johns MM, III, Ossoff RH. Burnout in academic chairs of otolaryngology: head and neck surgery. Laryngoscope 2005;115: 2056-2061.
19. Prins JT, Hoekstra-Weebers JE, Gazendam-Donofrio SM, Dillingh GS, Bakker AB, Huisman M, et al. Burnout and engagement among resident doctors in the Netherlands: a national study. Med Educ 2010;44: 236-247.
20. Saleh KJ, Quick JC, Conaway M, Sime WE, Martin W, Hurwitz S, et al. The prevalence and severity of burnout among academic orthopaedic departmental leaders. J Bone Joint Surg Am 2007;89: 896-903.
21. Tijdink JK, Smulders YM, Vergouwen ACM, de Vet HCW, Knol DL (2013) The assessment of publication pressure; validity and reliability of the publication pressure questionnaire (PPQ). Qual Life Res 2014; 23; 2055-2062
22. Streiner DL, Norman GR Cairney J. Health Measurement Scales : A practical guide to their development and use. 2008 Oxford University Press
23. Schaufeli WB, van Dierendonk D. Utrechtse Burnout Schaal (UBOS), handleiding [Utrecht Burnout Scale, manual]. Utrecht: 2000 Swets & Zeitlinger.
24. Maslach C, Jackson SE, Leiter MP. Maslach Burnout Inventory Manual. Palo Alto, CA: 1996. Consulting Psychologists Press.
25. Benedetti A, Abrahamowicz M. Using generalized additive models to reduce residual confounding. Stat Med 2004;23: 3781-3801.
26. Fanelli D. Do pressures to publish increase scientists' bias? An empirical support from US States Data. PLoS One 2010;5: e10271.
27. Anderson MS, Martinson BC, De Vries R. Normative dissonance in science: results from a notional survey of U.S. Scientists. J Empir Res Hum Res Ethics 2007; 2: 3-14.
28. Young NS, Ioannidis JPA, Al-Ubaydli O. Why current publication practices may distort science. PLoS Medicine 2008; 5: 201.
29. Shapin S. The scientific life: a moral history of a late modern vocation. 2008. Chicago: University of Chicago Press.
30. Davis MS, Riske-Morris, Diaz SR. Causal factors implicated in research misconduct: evidence from ORI case files. Sci Eng Ethics 2007;13: 395-414.
31. Sheldon T. Survey of Dutch doctors finds evidence of widespread research misconduct. BMJ 2012;344: 2898.
32. Louis KS, Anderson MS, Rosenberg L. Academic Misconduct and values - the departments influence. Review of higher education 2013;18: 393-422.
33. West CP, Huschka MM, Novotny PJ, Sloan JA, Kolars JC, et al. Association of perceived medical errors with resident distress and empathy: a prospective longitudinal study. JAMA 2006; 296: 1071-1078.

34. West CP, Tan AD, Habermann TM, Sloan JA, Shanafelt TD. Association of resident fatigue and distress with perceived medical errors. JAMA 2009;302: 1294-1300.
35. Cook C, Heath F, Thompson RL. A meta-analysis of response rates in web- or internet-based surveys. Educational and Psychological Measurement 2000;60: 821-836.
36. Klein DB, Stern C. Professors and their politics: the policy views of social scientists. Critical Review 2005;17: 257-303.
37. Joinson A. Social desirability, anonymity and Internet-based questionnaires. Behav Res Methods Instrum Comput 1999;31: 433-438.

5

Chapter 5

Publication Pressure and Scientific Misconduct in Medical Scientists

Joeri K. Tijdink, Reinout Verbeke, and Yvo M. Smulders

Journal of Empirical Research on Human Research Ethics 2014;9:64–71. doi: 10.1177/1556264614552421

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.

Introduction

Recent years have witnessed increased attention for scientific misconduct (1-3). 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" (4). For scientific misconduct observed in colleagues, corresponding rates were no less than 14% and 72%, respectively (5). 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, (6;7) 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 (6;8;9). Excessive emphasis on scientific output may also jeopardize academic activities that compete with science for time and attention, such as clinical and educational duties (10).

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 (11). 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.

Variables

The questionnaire contained, apart from demographic data, two parts: a validated Publication Pressure Questionnaire (PPQ) (9) 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" (12). 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 and Martinson (4;13). 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.

5

Table 1. Publication Pressure Questionnaire (PPQ).

Questions (domain)	Likert-type scale score (SD)	% agreement (4-5 or 1-2a 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.

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 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, 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.

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)

5

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

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.

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

	ß	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 1. Model 1: Demographic and job-specific characteristics; Model 2: Add adjustment for PPQ. PPQ = Publication Pressure Questionnaire.

Note 2. IN bold the variables with a significance level $p<0.05$.

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 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 (8). 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 (11). Insofar as this comparison is valid, publication pressure 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 (14;15). If we translate this to medical researchers, pressure and stress may contribute to scientific misconduct. It is conceivable that

Figure 1. Percentages of scientific misconduct.

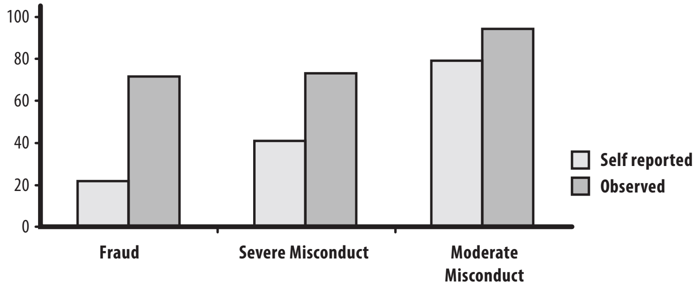
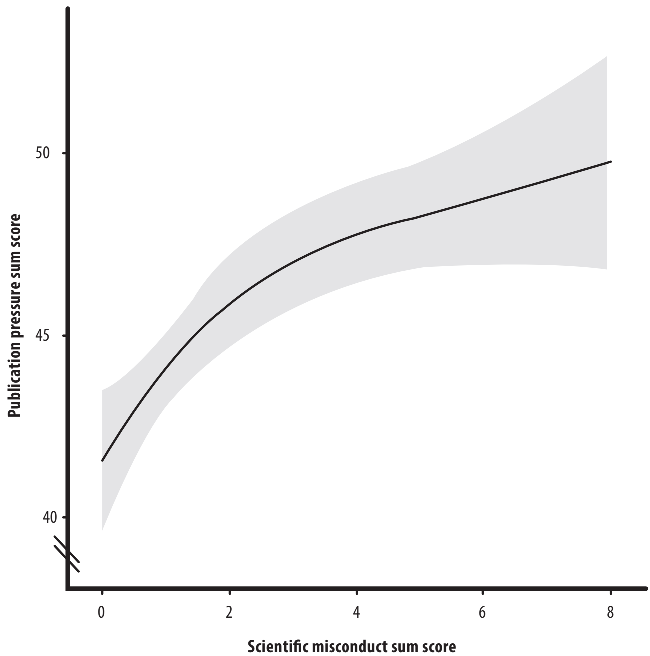


Figure 2. Cubic spline, demonstrating the association between the sum score on the validated 14-item Publication Pressure Questionnaire, and the scientific misconduct sum score.



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 (16).

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 (17-20). 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" (21). 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 (22).

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 (4;13). 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 (4). 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 (23;24). 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 (25).

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 (4). 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.

Reference List

(1) Couzin-Frankel J. Shaking up science. Science 2013;339:386-9.
(2) Fanelli D. Redefine misconduct as distorted reporting. Nature 2013;494:149.
(3) Wicherts JM. Psychology must learn a lesson from fraud case. Nature 2011;480:7.
(4) Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PLoS One 2009;4:e5738.
(5) Lafollette MC. The evolution of the "scientific misconduct" issue: An historical overview. Proceedings of the Society for Experimental Biology and Medicine 2000;224:211-5.
(6) Anderson MS, Ronning EA, DeVries R., Martinson BC. The perverse effects of competition on scientists' work and relationships. Science and Engineering Ethics 2007;13:437-61.
(7) Steneck NH. Fostering integrity in research: definitions, current knowledge, and future directions. Science and Engineering Ethics 2006;12:53-74.
(8) Fanelli D. Do pressures to publish increase scientists' bias? An empirical support from US States Data. PLoS One 2010;5:e10271.
(9) Tijdink JK, Smulders YM, Vergouwen AC, de Vet HC, Knol DL. The assessment of publication pressure in medical science; validity and reliability of a Publication Pressure Questionnaire (PPQ). Qualitiy of LIfe Research 2014;23:2055-62.
(10) Deangelis CD. Professors not professing. JAMA 2004;292:1060-1.
(11) Tijdink JK, Vergouwen AC, Smulders YM. Publication Pressure and Burn Out among Dutch Medical Professors: A Nationwide Survey. PLoS One 2013;8:e73381.
(12) Streiner DL, Norman GR. Health Measurement Scales : A practical guide to their development and use. Oxford University Press; 2008.
(13) Martinson BC, Anderson MS, DeVries R. Scientists behaving badly. Nature 20059;435:737-8.
(14) Pabst S, Brand M, Wolf OT. Stress effects on framed decisions: there are differences for gains and losses. Frontiers in Behavioral Neuroscience 2013;7:142.
(15) Vinkers CH, Zorn JV, Cornelisse S, Koot S, Houtepen LC, Olivier B, et al. Time-dependent changes in altruistic punishment following stress. Psychoneuroendocrinology 2013;38:1467-75.
(16) West CP, Tan AD, Habermann TM, Sloan JA, Shanafelt TD. Association of resident fatigue and distress with perceived medical errors. JAMA 2009;302:1294-300.
(17) Bedeian AG, Van Fleet DD, Hyman HH. Scientific achievement and editorial-board membership. Organizational Research Methods 2009;12:238.
(18) Bird SJ. Research ethics, research integrity and the responsible conduct of research. Science and Engineering Ethics 2006;12:412.
(19) Gannon F. The impact of the impact factor. EMBO Reports 2000;1:293.

(20) Hessels L.K. Science and the struggle for relevance Utrecht Univerity; 2010.

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

(22) Kalichman MW, Friedman PJ. A pilot study of biomedical trainees' perceptions concerning research ethics. Acadademic Medicine 1992;67:769-75.

(23) Bates SC, Cox JM. The impact of computer versus paper-pencil survey, and individual versus group administration, on self-reports of sensitive behaviours. Computers in Human Behaviour 2008;24:903-16.

(24) Tourangeau R, Smith T.W. Asking sensitive questions the impact of data collection mode, question format, and question context. Public opinion quarterly 1996;60:275-304.

(25) Joinson A. Social desirability, anonymity, and Internet-based questionnaires. Behavior Research Methods, Instruments, & Computers 1999;31:433-8.

Chapter 6

How do scientists perceive the current publication culture? A qualitative focus group interview study among Dutch biomedical researchers

6

J.K. Tijdink[1,6], K. Schipper[2], L.M. Bouter[3,4], P. Maclaine Pont[5], J. de Jonge[5], Y.M. Smulders[1]

BMJ Open 2016;6:e008681

Objective: To investigate the biomedical scientist's perception of the prevailing publication culture.

Design: Qualitative focus group interview study

Setting: 4 University Medical Centers in the Netherlands

Participants: 3 randomly selected groups of biomedical scientists (PhD, postdoctoral staff members and full professors)

Main outcome measures: Main themes for discussion were selected by participants.

Results: Frequently perceived detrimental effects of contemporary publication culture were the strong focus on citation measures (like the journal impact factor and the H-index), gift and ghost authorships and the order of authors, the peer review process, competition, the funding system, and publication bias. These themes were generally associated with detrimental and undesirable effects on publication practices and on the validity of reported results. Furthermore, senior scientists tended to display a more cynical perception of the publication culture than their junior colleagues. But even among the PhD students and the postdoctoral fellows the sentiment was quite negative. Positive perceptions of specific features of contemporary scientific and publication culture were rare.

Conclusions: Our findings suggest that the current publication culture leads to negative sentiments, counterproductive stress levels and - most importantly – to questionable research practices among both junior and senior biomedical scientists.

Background

The biomedical scientific enterprise has changed dramatically over the past decades. The annual number of published papers and scientific journals doubles every twelve years (1). There is an increasing imbalance between requested and available funding (2;3), raising concerns about hypercompetitiveness with potential distorting effects on the quality of research, the amount of research waste produced, the selection of priority research areas, and talent development (3-6). But some argue that increased demands on and competition between scientists has more beneficial than detrimental effects, and that a transparent reward system based on quantitative parameters is better than its alternatives (7). Regardless of how one evaluates these phenomena, the increasing emphasis on scientific productivity, authorships and citations by universities, grant agencies and indeed by the scientific community itself, is undeniable (8-10). The significant growth of the number of PhD dissertations puts an even greater pressure on the system (11). All the aforementioned phenomena are part of what can be described as 'publication culture'.

Earlier studies suggest that high publication pressure is associated with symptoms of burnout (12-15). Also, scientific integrity may be related to culture aspects of biomedical science (16-19).

Most of the aforementioned phenomena have been studied using quantitative survey methods, which provides some empirical basis for policy and future research, but may not capture all aspects and subtleties of scientists' views, thoughts and experiences. A qualitative approach, such as using focus group interviews, typically seeks to explore, understand and represent the subjective perceptions of people and to interpret their behaviour (20). This approach uncovers thoughts and feelings that survey research could never have highlighted and this has never been studied before. Focus group interviews are group conversations in which participants address specific themes (by sharing perspectives, experiences and opinions) (21).

We set out to perform focus group interviews about the perception of publication culture among PhD students, postdoctoral fellows/staff members and full professors that are involved in biomedical research. Our aim is to learn what biomedical scientists regard as the most salient aspects of current publication culture and to discuss the major positive and negative aspects of these features.

Materials and Methods

Selection of participants

The study consisted of 12 focus groups of biomedical scientists working in 4 university medical centres in The Netherlands. Scientists were eligible to participate if they were able to speak Dutch, were scientifically active (scientists that recently authored and published a scientific paper) and willing to give informed consent.

Scientists were recruited with the help of the deans' offices of the participating medical centres, that each provided e-mail addresses of all active scientists in 9 departments (2 pre-clinical (Microbiology, Pathology), 2 supportive departments (Methodology/Epidemiology, Anatomy/Physiology), 3 clinical departments (Internal Medicine, Surgery and Psychiatry), and the most and least publishing department (expressed as the average number of papers per active scientist). We used a tool specially designed by the software department of the VU university to randomly select the participants across the different academic ranks from the 9 departments. We randomly selected one PhD student, one postdoctoral fellow or staff member (usually an MD with a PhD degree, involved in a combination of patient care and research), and one full professor per department and per university medical centre, and sent an invitation per e-mail explaining the purpose of the focus group interviews. If the invited participant declined participation, we randomly selected a second participant of the same type from the same department, and so on, until we had 6-8 participants from different departments per focus group. This resulted in 3 focus groups (1 with PhD students, 1 with postdoctoral fellows and 1 with full professors) per university medical centre with 6-8 participating scientists per focus group.

Data collection and procedure

The focus groups were conducted between June 2013 and April 2014 by a multidisciplinary research team consisting of three of the authors of this article (JT, JdJ and PMP) at the 4 medical centres. Beforehand, the research team formulated possible discussion themes about publication culture based on our previous quantitative research on publication pressure (12) and a pilot version of a focus group interview that was conducted with fellow scientists from the department of the lead author. The focus group interviews lasted approximately 1.5-2 hours until the point when no new or relevant information emerged (attainment of saturation) (22;23).

The focus groups were led by a facilitator (JT and PMP) with professional experience in (focus) group dynamics. A semi-structured protocol (see supplementary material) was used, which included information on general aspects of focus groups, an introduction to the subject, and an initial exploration of the participants' motivation to be involved in research. After this, participants were invited to present themes they felt were relevant for the discussion on contemporary publication culture. From their answers, the facilitator, in consultation with the participants, prioritised 6-9 themes. Because it soon became clear that many comments and dominant opinions were negatively coloured, we explicitly encouraged participants to also name positive aspects of present publication culture.

Finally, participants were asked to suggest ways to solve the experienced problems (not part of this report).

Each focus group was audiotaped and transcribed verbatim. In addition, members of the research team took notes during the sessions to capture important elements.

Analysis

An inductive content analysis was used to analyse the data. Inductive content analysis is mainly used in cases where there are no or few previous studies dealing with the subject. A deductive approach is useful if the general aim is to test a previous theory in a different situation or to compare categories at different time periods (but that is clearly not at issue for our rarely studied topic) (24). By using an inductive content analysis we (JT, JdJ and PMP) read through the data looking for recurring themes. First, the entire transcripts were read and emerging themes were coded. New themes in the transcripts were added to the list of codes and added to the previously analysed results. The transcripts of the focus groups were analysed and coded independently by 3 team members (JT, PMP and JdJ) with different professional backgrounds (psychiatry, philosophy and sociology). Individual analyses of the team members were compared and discussed to achieve consensus and to increase reliability (25).

To check validity, participants received a written interpretation of the focus groups in which they participated, asking them to reflect (26) on our interpretation and to indicate if they recognised the analysis and coding. All participants agreed or had minor additional comments. Three team members (JT, JdJ and PMP) interpreted each of these transcripts and formulated the major themes discussed. This process of coding yielded 8 major themes. The results of the 12 focus groups were then compared, analysed, and interpreted by the 3 investigators, using an inductive approach. The final result was a summary of the 8 themes. Typical quotes were identified per theme and per scientific rank (PhD student, postdoctoral fellow/staff member and full professors) to clarify the coded themes. For review of the quality of reporting, the COREQ checklist was used.

Ethical considerations

All participants took part on a voluntary basis after giving consent by confirming participation through email. The study was not registered and reviewed by an ethics committee because the study only included scientists. Confidentiality was maintained using restricted, secure access to the data, destruction of audiotapes after transcription, and anonymous analysis of transcripts.

Inclusion of participants

We got 1810 email addresses of active scientists (stratified by department and by scientific rank) from 4 university medical centres (UMCs) in the Netherlands (UMC 1,2,3 and 4).

The 12 focus groups involved 79 participants (see table 1). The number of invitations that had to be sent out per included participant was: 1.75 for PhD students, 2.8 for postdoctoral staff members, and 2.16 for professors. Main reasons for declining participation were lack of time or having conflicting agendas.

Results

In the introduction of the focus group, participants were asked about their motivation for engaging in scientific research. Across all academic ranks and most strongly among PhDs, all participants most frequently reported curiosity and a quest for truth as their main driving force. Other less frequently described factors were to obtain funding and to show the world your scientific results. Among PhD students, an important motivation to start a PhD trajectory was to increase the chance of admittance to a residency programme for any of the medical specialties.

We identified 8 themes related to contemporary publication culture. Each theme is described below, and typical quotes that illustrate the opinions are reported in tables 2-4. Quotes in the tables are used as an illustration of the conclusions that were drawn per theme. Since the focus group interviews were conducted in Dutch, the quotes were translated into English by an official translation office for this report. The themes are presented in order of total frequency with which they were discussed in the 12 focus groups.

Research funding

A dominant perception across all focus groups was that there is hypercompetition for scarce funding. Furthermore, the procedures of funding agencies are generally perceived as being subjective and prone to manipulation, since participants felt that knowing the right people (committee members, reviewers of proposals) has a substantial impact on the chance of success.

To obtain funding, participants also mentioned the dominant role of the impact factors of journals in which publications were published, the number of publications and the Hirsch index. Finally, a common perception was that preparing grant applications was highly time consuming and thus expensive. Participants universally acknowledged that obtaining funding is a prerequisite for promotions and a bright career perspective.

Most participants believed that positive results are required to obtain funding. Comparing different focus group interviews in different academic ranks, it was obvious that for postdoctoral fellows and full professors funding is the most important. It can generate jobs and future job opportunities.

Authorships and author sequence

The second theme was authorships and author sequence. A frequently reported negative experience was that of disagreement regarding authorships and authorship sequence. According to the participants this is often an important cause for disputes in research groups.

This theme was also related to the importance of first and last authorships in the evaluation of institutions and individual scientists by funding agencies and universities. Also, most scientists considered the presence of gift authors (people who do not contribute significantly to a manuscript but are named in the author list) as a nuisance, even if it increases the chances of manuscript acceptance. 'If you don't contribute to a paper, you should not be on the author list'. Interestingly, rewarding team effort was hardly mentioned as a positive effect of the increased number of authors per paper.

A less frequently raised topic was the increasing number of authors. Some participants reported a sense of frustration, as multiple authors decreases reward and value of an authorship.

No positive comments were identified on this theme.

Quality versus quantity

The perceived tension between quality and quantity was a recurrent theme. Most scientists experienced individual performance evaluations as frustrating because the primary evaluation tool was felt to be the quantity of their scientific output rather than its quality. They expressed concerns regarding governmental policymakers who value journal impact factors more than scientific quality. Scientists also felt (albeit less often) that the number of publications is wrongly considered to be more important than societal impact or clinical relevance. Apart from these frustrations, professors and postdocs also perceived pressure to employ as many PhD students as possible, stimulated by the financial rewards for a doctorate. (In the Netherlands, government funding allocates a weight of 90,000 Euro to each PhD thesis.)

Some participants believed that the main motivation for biomedical PhD students to start a PhD trajectory was to improve their chance of obtaining a resident position in a medical specialty training programme. Such a lack of intrinsic scientific motivation could also affect scientific quality, according to focus group participants.

Except for occasional expressions of a sense of pride regarding Dutch 'publication efficiency' and number of publications per invested Euro, no positive comments were identified on this theme.

PhD students were more resentful that quantity is more important than quality in the present publication culture. This was of less concern to postdocs and professors.

Publication pressure

Although there is overlap between this theme and the theme quality versus quantity, participants identified publication pressure as a separate theme, mainly because publication pressure consists of more than quantity, it also includes the consequences for grant application success rates and position as a researcher. Publication pressure was also not directly linked to scientific quality by the participants.

Many focus group participants personally felt strong publication pressure, and had ideas about the underlying causes. They perceived a culture in which scientists are judged by the number of manuscripts published each year. Many they felt a strong pressure to obtain funding and to publish in high-impact journals in order to maintain their position in academia. This pressure was not only perceived as external, but also as a self-inflicted pressure.

Publication pressure was reported to compromise attention to other tasks, such as patient care or educational activities.

A minority of focus group participants experienced no publication pressure, but did notice such pressure among their colleagues.

Scientific misconduct and integrity

Scientists perceived ample room and opportunity to engage in questionable research practices (QRP). A commonly expressed cause for this was that research is perceived as solitary work: data analysis is often performed alone. There is little auditing by colleagues or fellow researchers, making scientific work vulnerable to QRP and research misconduct.

The participants also acknowledged that many biomedical scientists are not properly educated as to how to avoid the grey areas of QRP. Their perception was that much sloppy science is in fact due to a lack of sound methodological education, generating room for a grey area between responsible conduct of research and scientific fraud. According to virtually all participants, there is, in most cases, no intention to deceive readers.

All focus group participants felt that the pressure to publish positive results often stimulates scientist to cross the line of responsible conduct of research.

The PhD students reported that, due to strict hierarchy, they are reluctant to bring up QRP and research misconduct with their supervisors; they experience a lack of trust and confidentiality to talk to a senior researcher about possible research misconduct.

Many participants - especially postdocs and professors - expressed they can understand to some extent why some colleagues cannot resist the temptation of engaging in questionable research practices or even research misconduct.

A positive comment was that most participants thought scientific fraud is very rare in their communities; they felt there is almost never an intention to deceive.

Publication and reporting bias

Most participants felt that there is hardly any possibility to publish negative or 'no difference' results. For this they hold responsible the scientists, reviewers, editors and other stakeholders that take part in the publication process.

Many participants thought that 'sexiness of research results' (i.e. popular research areas with spectacular findings), rather than scientific quality, is essential to achieve high-impact publications.

Some participants expressed severe doubt as to whether high-impact journals judge and select submissions objectively based on scientific quality only, or whether they also select based on sexy results or citability.

Most scientists were aware that it is tempting to exaggerate their research results as a consequence of this 'positivitis'. As one associate professor said: 'The sexier the research results, the easier it gets published'.

As a consequence of published results being skewed towards positive outcomes, these results become difficult to replicate, according to many participants. No positive comments were identified on this theme.

Impact factor

Participants reported that, when they have to decide which journal to publish in, the Journal Impact Factor (JIF) is more important than the aim and scope of a journal. They felt, however, that judging a journal solely on its impact factor is wrong. Most participants emphasised that the impact factor is not a good index to measure scientific quality, as it predominantly measures impact based on recent citation scores, and does not necessarily reflect methodological rigor, let alone clinical relevance.

Some participants would not publish in journals with an impact factor <2 as they believed this could negatively impact their career. One professor felt he would be sanctioned by his superiors if he would publish in low impact-factor journals, because of effects on the ranking and prestige of his university.

Many PhD students expressed their outright anger about the extreme focus on the impact factor. They felt this was damaging to the scientific enterprise. Such frustrations were not expressed by more senior scientists.

A positively perceived aspect of the impact factor was that, although it is not a good indicator, it can to some extent help when deciding where to publish your research.

Disputes, prestige, self-satisfaction and competition

Many scientists experienced disputes among colleagues working in the same department. They believed that this is often caused by disagreements about authorships, envy and the unwillingness of some researchers to cooperate. Many also felt that scientists begrudge their colleagues' scientific success. Some participants believed that resentment and envy could negatively influence the quality of scientific studies, compromise peer review, and frustrate collaboration. Recognition and prestige were considered to be important personal factors in this process. As one professor stated: 'scientists can get high on a high-impact publication'.

Recognition and prestige were perceived as problematic mostly by experienced postdocs and professors. PhD students did not perceive this to be a major problem but emphasised the problem that sometimes they become involved, (to some degree) involuntarily, in disputes among senior researchers in their department.

A few participants also underlined the beneficial effects of competitiveness. They see competition as an essential ingredient for a flourishing, productive scientific culture.

Table 1. Dividing 79 participants among 12 focus groups. Between brackets the number of women.

	PhD students	Postdoctoral fellow/ staff member	Full Professors	Total
UMC 1	6 (3)	7 (3)	6 (2)	19 (8)
UMC 2	8 (3)	7 (4)	4 (0)	19 ((7)
UMC 3	8 (5)	3 (1)	6 (1)	17 (7)
UMC 4	8 (4)	8 (4)	8 (0)	24 (8)
Total	30 (15)	25 (12)	24 (3)	79 (30)

Table 2. Quotations to illustrate the content of the publication culture themes in PhD students

Theme	Quote
Research funding	‘you get grants because of friends and luck. Grants are no measure of ability but of who is who, who do you know and how you present it’
Authorships	‘Oh, we need to add that professor to the list of authors, because if he is on the list it will be easier to get accepted by such and such journal’
Quantity vs. Quality	‘What they measure now is how much and where you publish, but that says nothing about your qualities as a scientist’
Publication pressure	‘If the pressure on the number of publications decreases, the quality of the publications will increase’
Scientific integrity	‘(with regard to scientific integrity) It is not very common that the voice of the PhD student supersedes the voice of the person who is hierarchically superior. The boss calls the shots’
Publication Bias	‘If you find a positive association it is much easier to get published than in case of a negative result’
Impact factor	‘When you have an article published, the first question always is, what’s the impact factor. And if it is not very high they generally react; oh, but it is a really nice journal’
Competition, prestige, self-satisfaction and vanity	‘The loudest voice generally gets the best results’ ‘People often begrudge the other person also having his name on a paper’

Differences between scientific ranks

Most PhD candidates have rather naïve opinions about contemporary publication culture. They argue that science should be a genuine quest for truth and see scientists as truth-seekers who focus on scientific quality. Anything that disrupts this perception is judged negatively. Especially the present focus on the quantity of scientific output instead of scientific quality is a thorn in their side.

Postdoctoral fellows/staff members and professors hold more realistically or perhaps even slightly cynical views about the publication culture and are more sympathetic to the somewhat dubious elements in the scientific process. They accept these influences more readily.

Regarding publication pressure, the focus group interviews show that postdoctoral fellows/staff members feel the strongest pressure to publish. They experience the urge to produce in order to secure their positions and get the prestige and recognition for their publications, to get funded and prosper in their career (with a tenured professorship on the horizon). The present credit cycle in biomedicine mainly focuses on first authorship papers for PhD students and last authorships for professors. Postdoctoral fellows feel that they were sometimes denied last authorships, which in their opinion they deserved because of their role in the research process.

PhD students do not feel this amount of pressure, unless they are at the end of their PhD trajectory. Professors perceive less pressure than postdocs, since they already have a successful career and plenty of recognition.

Table 3. Quotations to illustrate the content of the publication culture themes in postdoctoral fellows/staff members

Theme	Quote
Research funding	'If you have no decent publications to put on your CV, you basically have no chance on the grants-market, that is what they look at, that is your fundraising capacity'
Authorships	'Authorship is a political game, sometimes you list someone as a co-author because you have to and you don't want an argument over something as trivial as one publication' 'If you confront him about it my boss becomes really angry and so I just list him' 'You often need a hotshot to be published in a high-impact journal. He often has to be the last author'
Quantity vs. Quality	'A lot of what is published is nonsense'
Publication pressure	'The stress of having to have at least 4-6 interesting and solid high-impact papers published each year; failure to produce means you will be judged to some extent'
Scientific integrity	'One is easily inclined to leave things out just to get it published'
Publication Bias	'That (publication bias) is the reason that fraud exists because without positive results I can forget about my career'
Impact factor	'That is what a professor said, that he preferred not to publish in lower-impact journals because it wouldn't look good on his CV'
Competition, prestige, self-satisfaction and vanity	'I think it is a universal quality of scientists that they are vain people, especially when they start publishing, they are often people who like the limelight and to be admired'

Table 4. Quotations to illustrate the content of the publication culture themes in full professors

Theme	Quote
Research funding	'The willingness to take risks continues to decrease whereas I feel that scientists should be willing to take risks, you see this especially when grants are involved
Authorships	'If you didn't feel so much pressure to publish you would also think more often that you don't need to have your name on a paper'
Quantity vs. Quality	'The highest goal of a professor is to deliver as many PhDs as possible, something I disagree with, by the way'
Publication pressure	'Too many publications are premature and slipshod'
Scientific integrity	'I think fraud and the pressure to publish are communicating vessels'
Publication Bias	'People want to be absolute, so everything (in papers, red.) is described in such a way that the message is earth shattering and unique; I get so tired of that'
Impact factor	'The scientific system, especially the biomedical disciplines, is totally fixated on impact factors, it's like a religion, when it's actually outdated'
Competition, prestige, self-satisfaction and vanity	'We need to be careful we don't get bogged down in measurements and who is the best' 'Publishing becomes such an idée fixe, such an important part of you … because you are published you are suddenly the man and then you may start to think you are a very important person'

Discussion

The purpose of our study was to identify and understand the perception of contemporary publication culture among Dutch biomedical scientists. Participants of the focus groups identified 8 themes in contemporary publication culture as relevant for their daily work.

In general, the current publication culture has a negative connotation, which is apparent in all 8 themes. With respect to research funding, participants expressed concerns over excessive competitiveness, unfairness, and lack of accessibility for newcomers and original concepts. Authorships and author sequence were commonly associated with disputes and conflicts among colleagues. Concerning quality versus quantity it was generally felt that the focus on the quantity of scientific output affected scientific quality. Publication pressure was not only described as an external source of stress from funding agencies and institutions, but also as an internal urge to improve personal career perspectives.

Engagement in questionable research practices and even in research misconduct was associated with pressure to publish, and participants did to some extent understand colleagues who could not resist the temptation to stray from a course of responsible conduct of research. The participants also believed that preferentially publishing positive findings (publication bias and positive outcome bias) in high-impact journals substantially improves scientific career perspectives. The impact factor (IF) has become increasingly dominant in current publication culture. Although the IF is not perceived as a quality predictor, it dominates the publication process. Participants regard the IF as one of the most important factors in deciding which journal they want to publish in.

Finally, the participants underline the important role of competition, prestige and vanity in scientists' motivation and conduct.

Comparison with existing literature

A previous focus group study among biomedical scientists in the USA (27) that investigated the role of competition in scientific practices found that competition has profound effects on the way science is performed. In that study, competitive experiences (such as prestige, grant application and pressure to publish) were perceived as detrimental and related to scientific integrity and personal job satisfaction. The results related to the theme competition are in line with these results.

Other research also supports the existence of a predominantly negative perception of publication culture. For example, competitiveness and a focus on productivity and citations have been related to perceived publication pressure (12). Excessive competitiveness is believed to have potentially perverse effects (27). Authors are reported to rush into print, cut corners, exaggerate their findings and overstate the importance of their results (3). These findings are confirmed by our participants.

The possible effects of a hypercompetitive scientific environment on scientific integrity are visible in frequent anecdotal reports (28). There is also empirical evidence in line with our findings; scientists who perceive high levels of pressure are more likely to withhold data or results (17;19) and studies suggest a correlation between the level of perceived competition, publication pressure, observed misconduct, fears of retaliation and conflicts (16;18;29). Nevertheless, the focus on publication culture as in the present study has never been systematically investigated.

Interpretation of the results

Our study addresses contemporary publication practices as seen through the eyes of biomedical scientists. But what do the results mean for the biomedical scientific community? Our results suggest that perceptions of current publication culture are mostly negative, causing a pessimistic and sometimes cynical view on (the validity of) scientific research.

Analysis of differences between job titles suggests that younger scientists hold a stronger view of science as a genuine quest for truth than many of their senior colleagues. Could this indicate a gradual decline of ideals over the course of a scientific career, caused by hypercompetitiveness? Or is the explanation found in the idealistic scientists preferring other career paths and leaving academia, causing selection of scientists as they become more senior? An answer can be found in the Cognitive Dissonance Theory (CDT) (30). Cognitive dissonance would mean that researchers who find themselves vested in a path that does not align with their ideals – hence in a state of conflicting attitudes, or cognitive dissonance – can either modify their behaviour (or quit) or modify their attitudes. The observed variation is congruent with the extent to which careers depend on publication pressure. Our study cannot differentiate between these and other possible explanations, but the finding itself calls for further research.

Limitations

Qualitative methods can be helpful when investigating complex, new or under-researched topics to generate hypotheses for further investigation (31). However, such studies lack advantages of quantitative studies, such as precise measures of effect sizes and variation.

Moreover, group dynamics can lead to distorted perspectives. The idea behind the focus group method is that group processes can help people to explore and clarify their views in group discussions with peers. On the one hand these dynamics may have caused exaggeration of some themes if an atmosphere of complaining and negativity in discussions develops in a group. On the other hand, group dynamics may have caused shyness to openly express every opinion, doubt or experience, thus causing underreporting and underestimation of themes, experiences and perceptions. Group work can actively facilitate the discussion of taboo topics because the less inhibited members of the group break the ice for shyer participants (32). This atmosphere was often created in the focus groups by our discussion leaders who have extensive experience with group dynamics.

Another factor that may have caused bias is prejudice in the group facilitators. Indeed the facilitators were part of research groups or organisations involved in assessment of research culture, and concerns over some aspects of research culture are indeed part of their everyday work. Nonetheless, facilitators with strong prejudices regarding likely outcomes could not guide focus groups, and instructions to facilitators were to be as objective as possible. They were instructed not to participate in discussions and to make sure that the participants decided for themselves which subjects and themes were discussed.

Regarding gender aspects, males were overrepresented in the full professor group. This is in accordance with the male/female ratio among professors in the Netherlands (12). Gender differences should be interpreted with caution in qualitative analysis. The study population was too small to draw firm gender related conclusions.

Considering the generalisability of the results, the sample is large enough to draw conclusions. The results can be seen as reasonably valid, as we reached saturation per layer (22;23). Nevertheless, the reader must decide, interpret and reflect whether the results are generalizable for their scientific practice. It can be questioned whether our findings apply to other countries. Academic structure and culture in other countries may certainly differ. Nevertheless, the problems that were presented in the focus group study by Melissa Anderson (27) showed similar results in the US. Furthermore, publication pressure measured by the Publication Pressure Questionnaire (PPQ) was also high in a Flemish population (18).

Finally, the influence of a response bias cannot be ruled out. The number of invitations that had to be sent per participant was 1.75 for PhD students, 2.8 for postdoctoral staff members, and 2.16 for professors. Most invited scientists who did not participate were asked to explain their reasons for declining participation. Reasons such as lack of time, conflicting agendas, maternity leaves or no mastery of Dutch language were mentioned. Nevertheless, we cannot exclude unwillingness to participate as a possible source of response bias.

Changing the culture

It is not easy to push an established culture in another direction. Academic structure is complex, which makes it hard to predict which interventions will work and to whom they should be directed. Nevertheless, change starts with increased awareness among all parties involved. In this light, the good news is that numerous initiatives across different scientific areas have recently emerged.

(To name a few: METRICS, the DORA manifesto, Force11, ALTmetrics, Science in Transition, the REWARD alliance etc.) These initiatives will eventually result in new values and forms to reshape current publication practices.

Conclusion

Active biomedical scientists from 4 university medical centres in the Netherlands describe a publication culture with an extreme focus on impact factors, funding, authorships and publishing positive papers. These factors intensify competition between them and emphasise the dominance of quantitative scientific output over methodological quality, especially over the replicability of findings. This raises serious concerns about the credibility of scientific results. Future research should identify alternatives and interventions to restore core values such as trust, credibility, integrity and collaboration.

Reference List

(1) Ridker PM, Rifai N. Expanding options for scientific publication: is more always better? Circulation 2013;127(2):155-6.

(2) Freeman R, Weinstein E, Marincola E, Rosenbaum J, Solomon F. Careers. Competition and careers in biosciences. Science 2001;294(5550):2293-4.

(3) Alberts B, Kirschner MW, Tilghman S, Varmus H. Rescuing US biomedical research from its systemic flaws. Proc Natl Acad Sci U S A 2014;111(16):5773-7.

(4) Al-Shahi SR, Beller E, Kagan J, Hemminki E, Phillips RS, Savulescu J, et al. Increasing value and reducing waste in biomedical research regulation and management. Lancet 2014;383(9912):176-85.

(5) Ioannidis JP, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet 2014;383(9912):166-75.

(6) Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP, et al. Biomedical research: increasing value, reducing waste. Lancet 2014;383(9912):101-4.

(7) Thomas L.G.III. The two faces of competition: Dynamic resourcefulness and the hypercompetitive shift. Organization Science 1996;7:221-42.

(8) Hessels L.K. Science and the struggle for relevance. Utrecht Univerity; 2010.

(9) Ioannidis JP, Boyack KW, Small H, Sorensen AA, Klavans R. Bibliometrics: Is your most cited work your best? Nature 2014;514(7524):561-2.

(10) Ioannidis JP, Khoury MJ. Assessing value in biomedical research: the PQRST of appraisal and reward. JAMA 2014;312(5):483-4.

(11) Cyranoski D, Gilbert N, Ledford H, Nayar A, Yahia M. Education: The PhD factory. Nature 2011;472(7343):276-9.

(12) Tijdink JK, Vergouwen AC, Smulders YM. Publication Pressure and Burn Out among Dutch Medical Professors: A Nationwide Survey. PLoS One 2013;8(9):e73381.

(13) Tijdink JK, Vergouwen AC, Smulders YM. Emotional exhaustion and burnout among medical professors; a nationwide survey. BMC Med Educ 2014;14(1):183.

(14) Miller AN, Taylor SG, Bedeian AG. Publish or perish: academic life as management faculty live it. Career development international 2011;16(5):422-45.

(15) van Dalen HP, Henkens K. Intended and unintended consequences of a publish-or-perish culture: A worldwide survey. Journal of the American Society for Information Science and Technology 2012;63(7):1282-93.

(16) Anderson M.S. Misconduct and departmental context: Evidence from the Acadia Institute's graduate education project. Journal of Information Ethics 1996;5(1):15-33.

(17) Blumenthal D, Campbell EG, Gokhale M, Yucel R, Clarridge B, Hilgartner S, et al. Data withholding in genetics and the other life sciences: prevalences and predictors. Acad Med 2006;81(2):137-45.

(18) Tijdink JK, Verbeke R, Smulders YM. Publication pressure and scientific misconduct in medical scientists. Journal of Empirical Research and Human Research Ethics 2014;9(5):64-71.
(19) Walsh JP, Hong W. Secrecy is increasing in step with competition. Nature 2003;422(6934):801-2.
(20) Creswell J.W. Qualitative inquiry and research design: choosing among five traditions. Thousand Oaks: Sage; 1998.
(21) Krueger R.A., Casey M.A. Focus groups: A Practical Guide for Applied Research. 3rd ed. Thousand Oaks, CA. Sage Publications; 2000.
(22) Barbour RS. Checklists for improving rigour in qualitative research: a case of the tail wagging the dog? BMJ 2001;322(7294):1115-7.
(23) Mays N, Pope C. Qualitative research in health care. Assessing quality in qualitative research. BMJ 2000;320(7226):50-2.
(24) Elo S, Kyngas H. The qualitative content analysis process. J Adv Nurs 2008;62(1) 107-115.
(25) Lincoln Y.S., Guba E.G. Naturalistic Inquiry. Newbury Park, CA: Sage Publications; 1985.
(26) Meadows L.M., Morse J.M. Constructing evidence within the qualitative project. In: Morse J.M., Swanson J.M.& Kuzel A.J., editors. The Nature of qualitative evidence. Thousand Oaks, CA: Sage publications; 2001. p. 187-200.
(27) Anderson MS, Ronning EA, DeVries R., Martinson BC. The perverse effects of competition on scientists' work and relationships. Science and Engineering Ethics 2007;13(4):437-61.
(28) Kennedy D. Academic Duty. Cambridge, MA: Harvard University Press; 1997.
(29) Louis K.S., Anderson M.S., Rosenberg L. Academic misconduct and values: The department's influence. The Review of Higher Education 1995;18:393-422.
(30) Festinger L. A Theory of Cognitive Dissonance. Stanford, California: Row, Peterson and Company; 1957.
(31) Mays N, Pope C. Rigour and qualitative research. BMJ 1995;311(6997):109-12.
(32) Kitzinger J. Qualitative research. Introducing focus groups. BMJ 1995;311(7000): 299-302.

7

Chapter 7

How does industry funding disclosure influence psychiatrists? A randomized trial among Dutch psychiatrists

7

Joeri K Tijdink, Yvo M Smulders, Lex M Bouter, Christiaan H Vinkers

Under review 2016

Abstract

Background: It has been shown that most industry sponsored studies report positive results rather than negative or inconclusive results. It is unknown whether mentioning of study funding source influences the appraisal of an RCT by psychiatrists. Additionally, it is unclear whether reporting positive findings influences perceived credibility or clinical relevance of a study. This study investigates whether psychiatrists' appraisal of a scientific abstract is influenced by industry funding disclosures and/or by positive outcomes.

Methods: In this study, Dutch psychiatrists were randomized to evaluate a scientific abstract describing a fictional RCT for a novel antipsychotic drug. Four different abstracts were created reporting either absence or presence of industry funding disclosure as well as a positive or a negative outcome. Primary outcomes were the perceived credibility and clinical relevance of the study results (10-point Likert scale). Secondary outcomes were the assessment of methodological quality and interest in reading the full article.

Results: 395 psychiatrists completed the survey (completion rate45 %). Industry funding disclosure was found not to influence perceived credibility (MD 0.12; 95% CI -0.28 to 0.47) nor sense of clinical relevance of the study (MD 0.14; 95% CI -0.54 to 0.27). A negative outcome was perceived as more credible than a positive outcome (Mean Difference (MD) 0.81 points; 95% Confidence Interval (CI) 0.43 to 1.18), but generated similar scores for clinical relevance (MD -0.14; 95% CI -0.54 to 0.27).

Conclusions: In this study, industry funding disclosure did not affect the perceived credibility nor judgement of clinical relevance. Psychiatrists perceive positive study outcomes to be less credible than negative outcomes. These results suggest that psychiatrists fail to recognize the influence of funding sources on the research results. This should be taken into account in the ongoing debate on industry-physician relationships.

Introduction

Several factors are known to influence the interpretation of research findings by editors, reviewers, readers and fellow scientists (1-4). Of these factors, study outcome and funding disclosure are important (2;5). The excess of positive results in the scientific literature, most likely due to selective reporting, is widely acknowledged (6;7). In addition, it has been firmly established that industry sponsored studies more frequently report positive results when compared to non-sponsored trials (8;9), and that many 'negative' industry-sponsored studies remain unpublished (10-13). However, there is a limited number of studies on the factors that influence physicians in their interpretation of the scientific literature. This topic is important since physicians often fail to recognize the impact of conflicts of interest (14), even though a critical attitude by medical doctors towards industry funding has been found to affect the perceived relevance of a study (15). In this context, psychiatry is a field of particular interest. Psychiatrists are frequently criticized for their ties to the pharmaceutical industry and the impact of these ties has been intensely debated by top flight journals (16-18). Pharmaceutical companies have been criticized for the high profits they received from antidepressants and antipsychotic sells. Their sells may have been influenced by substantial publication bias (13). Furthermore, a controversy has emerged from reanalysis of data from sponsored study 329 that showed serious side effects of antidepressants which were not reported in the initial study (19). Nevertheless, it is unknown whether these and other factors such as study outcomes and industry funding influence the scientific evaluation of an RCT. If psychiatrists are easily swayed by funding disclosure or positive outcomes, this could have direct impact on their prescribing behavior and influence their clinical decision making. Therefore, we aimed to assess how study outcome and industry funding disclosure influence psychiatrists' perception of credibility and clinical relevance of results of a hypothetical RCT.

Methods

General

The ethical review board of the VU Medical Center considered the research project and decided that no formal ethical approval was necessary for this study.

Abstract development

Three of the authors of this study extensively discussed the content of the proposed fictitious abstracts. We used the PubMed format (reproduced without the NCBI logo in supplementary file SF1-4) to make it appear like an original study. After producing a first draft of the abstracts by three authors, it was sent as a pilot version to 5 other psychiatrists to receive feedback on formulation, design, credibility and face-validity. This feedback resulted in minor modifications and thus the final version of the abstract. Each participant received the survey in Dutch containing an abstract in the English language describing an RCT for a non-existing novel antipsychotic drug 'vinquerine' in a fictitious journal ("the Archives of Clinical Psychiatry"). We chose to display the abstract in English since Dutch psychiatrists are comfortable in reading scientific literature in English. Four different abstracts were created based on reported study outcome (positive (a statistical significant difference compared with olanzapine) vs negative (no statistical differences compared with olanzapine)) and industry funding disclosure (yes/no). The original abstracts are included in the supplementary material (Supplementary Figures in SF1-4 Figure). The fictional study compared the effects of vinquerine to those of olanzapine and placebo on psychotic symptoms in patients with a first-episode psychosis. We chose olanzapine because it is an established and often used treatment for psychosis. Severe side effects were recorded in the fictitious abstract since these influence prescription behavior; the presence or absence of side effects often makes significant influence on the choices clinicians make in the treatment of psychiatric patients.

A positive outcome was reported by showing a statistically and clinically significant effect of vinquerine compared to olanzapine (an often applied first line treatment for psychosis) and placebo. In the positive outcome abstract, vinquerine was reported to have very limited side effects.

A negative outcome was reported as vinquerine and olanzapine being equally effective, but vinquerine showing important side effects. Both vinquerine and olanzapine were reported as being superior to placebo.

In the industry funded version, the second author was reported to be a consultant for a fictional pharmaceutical company ('Olevy Pharmaceuticals'). We chose to put the COI under the second author, rather than the first, since it was considered to be more consistent with existing scientific reporting. At the bottom of the fictitious industry funded abstracts, it was clearly stated that 'this study was funded by a research grant from Olevy Pharmaceuticals'.

Identical methodological limitations were deliberately introduced in all versions of the abstract to make it consistent with Pubmed abstracts that, according to the authors, were judged to be average quality. We felt that reporting methods of average quality would make it more likely that the article would be scrutinized for validity. In contrast, a study without noticeable limitations could be received by clinicians as being fabricated or lacking veracity.

The limitations included a relatively small sample size (n=303), unclear selection of study participants, non-blinded study design, a relatively short follow-up period (4-weeks), and the exclusion of non-compliant patients in a per-protocol analysis. The abstract-only approach was deliberately chosen since physicians frequently only read the abstract of potentially interesting scientific studies (20).

Survey sample and randomization

VanDerHoef&Partners provided to the authors the email addresses of Dutch psychiatrists but were not involved in any part of the study (concept, design, analysis, and writing). 1566 Dutch psychiatrists were randomized to receive an e-mail with a link to one of the following four abstracts: 1) negative outcome/industry funding disclosure, 2) negative outcome/no industry funding disclosure, 3) positive outcome/industry funding disclosure, or 4) positive outcome/no industry funding disclosure. The invitational e-mail and subsequent reminders were sent in May and June 2014 and included a brief statement to describe the goal of the study in very general terms ('to determine how psychiatrists evaluate scientific research'). Participants were left unaware of the true design and intention of the study. The e-mail included a link to the online survey and the instructions on how-to opt-out of this study if they so chose. Two reminders were sent within a four week time frame to psychiatrists who did not respond to the initial email. If psychiatrists declined to respond, they were asked to follow an electronic link to a very brief online questionnaire to disclose their reasons for non-participation.

Survey characteristics and outcomes

Each abstract was accompanied by a short questionnaire with 10 statements to determine the credibility ("How would you rate the credibility of the abstract?"), clinical relevance, methodological rigor (7 statements), and interest in reading the full article using a 10-points Likert scale (1: very poor, 10; extremely good). A control question ("please enter the score 3 for this question") was included to check if participants did not randomly complete the survey. Methodological rigor was assessed with seven items, addressing different methodological characteristics: study design, methodology, statistical analysis, sample size, outcome measures, completeness of reporting and overall study quality [see supplementary table in S1 Table]. An equal-weighted sum score for the 7 items was calculated. To prevent that the order of questions have an undue influence on the answers respondents provided, the primary outcome questions were randomly distributed among the 10 questions.

After completion of the survey, participants' attitudes towards the pharmaceutical industry were assessed. We also asked for financial ties with the pharmaceutical industry. Participants were asked if they had received a representative of a pharmaceutical company or if they had received any industry funding in the past 6 months. In addition, theyr were provided with four statements regarding the influence of industry funding on scientific clinical studies. Participants were asked to agree or disagree. At this stage of the survey, participants were not able to alter any of their previous answers.

Participant characteristics were obtained by asking respondents to report their gender, age, professional affiliations (academia, general hospital, or mental health care facility) and whether or not they had obtained a PhD. In addition, they were asked their place of residency, and whether the participant was actively involved in scientific research.

Statistical analysis

Primary outcome measures were the perceived credibility and the judgement regarding clinical relevance of the study results reported in the fictitious abstract. Secondary outcomes were the level of interest in reading the full article, and global assessment of methodological rigor (sum score). All primary and secondary outcomes were scored on a 10-point Likert scale. The survey software was constructed in a manner that prevented participants from submitting the survey unless all questions were completed. This was done to minimize missing data. First, possible interaction between the industry funding disclosure and study outcome was addressed. For the primary outcomes, the effect of funding disclosure and study outcome were analyzed using 2x2 ANOVA. Possible confounders and effect modifiers were analyzed, including participants' self-reported attitude towards industry funding, and active relationships with the pharmaceutical industry. As a secondary, exploratory analysis, linear univariate and multivariate regression analyses were used to identify other possible determinants of perceived credibility and clinical relevance, including demographic and job specific factors, and active relations with industry. The Statistical Package for the Social Sciences (SPSS) statistics (Chicago USA 2011, version 20) was used for all statistical analyses.

Results

Participant Characteristics

A total of 880 psychiatrists opened the invitational email, to which 580 psychiatrists responded (66%) and 395 (45%) completed the full survey. Two respondents were excluded because they failed to correctly answer the control question. Demographic data and characteristics of the participants are presented in Table 1. Attitudes towards industry funding are included in supplementary tables in S2&S3 tables.

Industry funding disclosure and study outcome

Interaction analysis

No interaction was found between study outcomes and industry funding disclosure on the primary outcomes for both 'credibility' ($p=0.99$) and 'clinical relevance' ($p=0.41$).

Table 1. Demographic and professional characteristics of the participants.

		N=395	%
Gender	Male (%)	214	54
Age	Mean (range)	50 (27-72)	
Professional affiliation	Academia	40	10
	General Hospital	37	9
	Mental Health Institution	248	63
	Private Practice	63	16
	Other	38	10
Residency background	Academic Hospital	158	40
	General Hospital	4	1
	Mental Health Institution	200	51
	Other	32	8
Sub-specialty	Adult psychiatry	241	61
	Child psychiatry	90	23
	Geriatric psychiatry	25	6
	Other	37	9
Received representative of a pharm. company	Yes	180	46
Received pharmacological funding	Yes	14	4
Scientifically active	Yes	131	33
PhD degree	Yes	104	27
Years employed as psychiatrist (range)		14 (0-40)	

Industry funding disclosure does not affect credibility or relevance assessments

Industry funding disclosure was not significantly associated with perceived credibility (MD 0.12; 95% CI -0.28 to 0.47) or clinical relevance (MD 0.14; 95% CI -0.54 to 0.27). Likewise, no significant effects were found for industry funding disclosure on the secondary outcomes 'assessment of methodological rigor' (MD 0.22; 95% CI -1.82 to 2.17), and 'interest in reading the full article' (MD 0.14; 95% CI -0.40 to 0.71) (Table 2).

Positive study outcomes are perceived as less credible

A positive study outcome of the fictional RCT was associated with significantly lower scores on credibility (MD 0.81; 95% CI 0.43 to 1.18) but not 'clinical relevance' (MD 0.14; 95% CI -0.28 to 0.53), compared to a negative study outcome. The secondary outcome 'interest in reading the full article' (MD 0.54;

Table 2. ANOVA analysis of the primary and secondary outcomes of the abstract with or without funding disclosure and with a positive or negative study outcome

	Funding (N=206)	No Funding (N=187)	Mean Difference (MD)	95% CI
Credibility #	4.64	4.76	0.12	-0.28 to 0.47
Clinical relevance	5.30	5.16	0.14	-0.54 to 0.27
Interest in reading the full article	4.51	4.65	0.14	-0.40 to 0.71
Methodological quality (sum score) ##	38.65	38.87	0.22	-1.82 to 2.17
	Negative outcome (N=200)	Positive outcome (N=193)	Mean Difference (MD)	95% CI
Credibility#	**5.10**	**4.29**	**0.81**	**0.43 to 1.18**
Clinical relevance#	5.16	5.30	0.14	-0.28 to 0.53
Interest in reading the full article #	**4.31**	**4.85**	**0.54**	**0.09 to 1.12**
Methodological quality (sum score)##	5,63	5,44	0.19	-3.31 to 0.68

10-point Likertscale score

Average score on a 10-point Likertscale of the seven individual items regarding methodological quality (see supplementary table S1).

95% CI 0.09 to 1.12) was statistically significant. The other outcome 'assessment of methodological rigor' (MD 0.19; 95% CI -3.31 to 0.68) was not significantly influenced by the reported study outcome (Table 2).

No effect modification by relations with or attitude towards industry funding

We investigated whether participants' active relations with- and attitude towards-industry were effect modifiers or confounders in the observed relations between study outcome and the primary outcomes. Both for perceived credibility and clinical relevance, no interactions or confounding were found for pharmaceutical industry relations and attitude towards industry funding (data not shown). Nevertheless, psychiatrists were well aware of possible industry effects, with an average score of 8.0 of 10 for the statement that "*a pharmaceutical company can influence study results*", a score of 7.4 out of 10 for the statement that "f*unding* *has an effect on the quality of research*", and 6.9 out of 10 for the statement that "*funding from a pharmaceutical company has a negative influence on the validity of research results*" (see suppl. table in S3 table).

Professional characteristics that influence perceived credibility and clinical relevance

In a secondary exploratory analysis, we investigated whether professional characteristics affect perceived credibility and clinical relevance scores in the total study population. Credibility of the RCT was negatively associated with having scientific experience (e.g. having a PhD or being scientifically active) (MD -0.6; 95% CI -0.99 to -0.21), and positively associated with having recently received a pharmaceutical representative (MD 0,63; 95% CI 0.25 to 1.01), funding from a pharmaceutical company (MD 1.27; 95% CI 0.25 to 2.30) or being employed in a general mental health institution (MD 0,46; CI 0,06 to 0,85). Multivariate analyses of these variables did not affect any of these results (data not shown). For participants' assessment of clinical relevance, psychiatrists interpreted this more negative if they had scientific experience (PhD or being scientifically active) (MD -0.43; 95% CI -0.85 to -0.02), were longer active as a psychiatrist (beta -0.02; 95% CI -0.04 to 0.00) and had received residency training in a general mental health institution rather than an academic hospital (MD 0.40; 95% CI -0.80 to 0.00). Multivariate analysis only showed significant associations for scientific experience and residency training in a general non-academic mental health institution (data not shown).

Discussion

This study demonstrates that psychiatrists have more confidence in the validity of a negative outcome than a positive outcome of a fictitious study assessing a novel antipsychotic drug. In apparent contrast, the evaluation of a scientific abstract appears insensitive to industry funding disclosure regardless of the reader's attitudes towards or relationships with the pharmaceutical industry. The relative distrust of a positive study outcome was demonstrated and is in agreement with the widespread recognition that the medical literature suffers from positive outcome bias (6;7), particularly in drug studies (21;22). Positive outcome bias has also been convincingly demonstrated in psychiatric literature, for example on the effectiveness of antidepressants (4;23) and antipsychotics (24). Presumably, respondents were aware of this bias, generating hesitation to accept results that may be considered 'too good to be true'. The fact that positive outcomes are more critically appraised by psychiatrists is consistent with another study that highlights the influence of positive outcomes in paper acceptance rates in peer review (2). These results are not fully comparable as peer reviewers have other criteria to assess a scientific paper than clinicians.

Although the detrimental effects of industry funding on study outcomes has been well-established both in the general medical literature (5) and in the psychiatric literature (11), we were surprised that no effect was found for funding disclosure on the credibility or perceived relevance of, in particular the positive outcome version of the abstract. Moreover, participants' attitudes towards pharmaceutical industry funding did not influence participants' appraisal of the fictitious abstract. (see table in S3 Table). Participants with overall high scores on statements that pharmaceutical companies can unduly influence study results were no more likely to question the validity of the fictitious abstract, even though the abstract clearly disclosed the funding. We can only speculate why these attitudes did not affect the results of our study. Apparently, psychiatrists did not automatically connect their perception of funding effects to the actual interpretation of the abstract we sent them.

It also has to be considered that disclosure of industry funding in itself does not make study results more or less valid (25).

The recent media attention on study 329 in the BMJ reinforces the need for individual patient data and original study protocols to increase the validity of scientific results as registration of clinical trials is insufficient (19).

Around 27% of the participants had a PhD degree, which is higher than the average among community psychiatrists. Probably, psychiatrists with a PhD are more likely to participate in our survey, as they would feel more competent to assess a scientific abstract.

Secondary analyses showed that credibility scores were higher in psychiatrists with active relations with pharmaceutical companies and this was independent of industry funding disclosure. Credibility scores of the fictitious abstract were lower for the psychiatrists with scientific research experience. These results suggest that participants' relationships with the pharmaceutical industry can influence them to appraise study results of scientific research more positive or less critical.

To our knowledge, this is the first randomized study to investigate whether industry funding disclosure and study outcomes can influence the opinion of psychiatrists on the credibility and relevance of clinical research. In contrast, a somewhat smaller, previous study among 269 internists suggested that industry funding disclosure resulted in lower scores on methodological rigor and less confidence in the results (15). The study design with different methodological approaches and the inclusion of three hypothetical drugs may have played a role in the discrepancy between these two studies. Additionally, it may be possible that internists are more critical towards pharmaceutical industry funded studies compared to psychiatrists.

The results of this study should be cautiously interpreted in light of several limitations. First, internet-based questionnaires can suffer from a selective response bias, predominantly attracting participants who feel capable of assessing a scientific abstract. In our study, the response rate was relatively high compared to other internet-based surveys for physicians (26;27). Furthermore, randomization of the participants should have minimized the effects of a selective response.

Secondly, it can be debated whether a 19% difference in perceived credibility score between the two groups is large enough to allow firm conclusions.

Thirdly, it might be that the participants perceived the positive study as invalid because olanzapine is a well-established reference antipsychotic drug. Furthermore, they might have perceived the negative study as more realistic in terms of symptom reduction. Finally, some respondents may not have noticed the disclosure of industry funding in the scientific abstract even though all seven psychiatrists who pretested the fictitious abstract had noticed this.

Conclusion

This study suggests that psychiatrists fail to account for the effects of industry funding disclosure in their judgement of the reliability and relevance of study results.. On the other hand, psychiatrists are more likely to critically interpret the content of a scientific abstract in which a positive outcome is reported. Our results are timely in light of the recent discussion on the effects of industry-physician relations in several leading medical journals (17;18;28). From our study, it is apparent that there is a striking discrepancy between psychiatrists' attitudes towards the pharmaceutical industry and the effects that funding disclosure has on their perceived credibility and judgement of clinical relevance instudy results. Studies addressing effects that positive outcomes and industry funding disclosures have on readers and on clinical decision making are critically important and should be taken into account in any discussions regarding the effects of these factors.

Acknowledgement

We would like to thank VanDerHoef&Partners (Locum tenens company) for providing email addresses of the participating psychiatrists.

We would like to thank Dr. James McKinnon, psychiatrist for his critical review of the manuscript in English.

Reference List

(1) Angell M. Industry-sponsored clinical research: a broken system. JAMA 2008;300: 1069-71.

(2) Emerson GB, Warme WJ, Wolf FM, Heckman JD, Brand RA, Leopold SS. Testing for the presence of positive-outcome bias in peer review: a randomized controlled trial. Arch Intern Med 2010;170:1934-9.

(3) Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. Cochrane Database Syst Rev 2012;12:MR000033.

(4) Turner EH, Matthews AM, Linardatos E, Tell RA, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. N Engl J Med 2008;358:252-60.

(5) Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. BMJ 2003;326:1167-70.

(6) Chan AW, Hrobjartsson A, Haahr MT, Gotzsche PC, Altman DG. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. JAMA 2004;291:2457-65.

(7) Dwan K, Gamble C, Williamson PR, Kirkham JJ. Systematic review of the empirical evidence of study publication bias and outcome reporting bias - an updated review. PLoS One 2013;8:e66844.

(8) Bhandari M, Busse JW, Jackowski D, Montori VM, Schunemann H, Sprague S, et al. Association between industry funding and statistically significant pro-industry findings in medical and surgical randomized trials. CMAJ 2004;170:477-80.

(9) Bourgeois FT, Murthy S, Mandl KD. Outcome reporting among drug trials registered in ClinicalTrials.gov. Ann Intern Med 2010;153:158-66.

(10) Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev 2009;(1):MR000006.

(11) Perlis RH, Perlis CS, Wu Y, Hwang C, Joseph M, Nierenberg AA. Industry sponsorship and financial conflict of interest in the reporting of clinical trials in psychiatry. Am J Psychiatry 2005;162:1957-60.

(12) Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: review of publication and presentation. PLoS Med 2008;5:e217.

(13) Turner EH. Publication bias, with a focus on psychiatry: causes and solutions. CNS Drugs 2013;27:457-68.

(14) Silverman GK, Loewenstein GF, Anderson BL, Ubel PA, Zinberg S, Schulkin J. Failure to discount for conflict of interest when evaluating medical literature: a randomised trial of physicians. J Med Ethics 2010;36:265-70.

(15) Kesselheim AS, Robertson CT, Myers JA, Rose SL, Gillet V, Ross KM, et al. A randomized study of how physicians interpret research funding disclosures. N Engl J Med 2012;367:1119-27.
(16) Appelbaum P., Arboleda-Florez J., Javed A., Soldatos C., Tyano S. WPA recommendations for relationships of psychiatrists, health care organizations working in the psychiatric field and psychiatric associations with the pharmaceutical industry. World Psychiatry 2011;10:155-9.
(17) Godlee F. Conflict of interest: forward not backward. BMJ 2015;350:h3176.
(18) Horton R. Offline: The BMJ vs NEJM--lessons for us all. The Lancet 2015;385:2238.
(19) Le Noury J, Nardo JM, Healy D, Jureidini J, Raven M, Tufanaru C, et al. Restoring Study 329: efficacy and harms of paroxetine and imipramine in treatment of major depression in adolescence. BMJ 2015;351:h4320.
(20) Saint S, Christakis DA, Saha S, Elmore JG, Welsh DE, Baker P, et al. Journal reading habits of internists. J Gen Intern Med 2000;15:881-4.
(21) Page MJ, McKenzie JE, Kirkham J, Dwan K, Kramer S, Green S, et al. Bias due to selective inclusion and reporting of outcomes and analyses in systematic reviews of randomised trials of healthcare interventions. Cochrane Database Syst Rev 2014;10:MR000035.
(22) Saini P, Loke YK, Gamble C, Altman DG, Williamson PR, Kirkham JJ. Selective reporting bias of harm outcomes within studies: findings from a cohort of systematic reviews. BMJ 2014;349:g6501.
(23) Roest AM, de Jonge P., Williams CD, de Vries YA, Schoevers RA, Turner EH. Reporting Bias in Clinical Trials Investigating the Efficacy of Second-Generation Antidepressants in the Treatment of Anxiety Disorders: A Report of 2 Meta-analyses. JAMA Psychiatry 2015;72:500-10.
(24) Turner EH, Knoepflmacher D, Shapley L. Publication bias in antipsychotic trials: an analysis of efficacy comparing the published literature to the US Food and Drug Administration database. PLoS Med 2012;9:e1001189.
(25) Loewenstein G, Sah S, Cain DM. The unintended consequences of conflict of interest disclosure. JAMA 2012;307:669-70.
(26) Asch DA, Jedrziewski MK, Christakis NA. Response rates to mail surveys published in medical journals. J Clin Epidemiol 1997;50:1129-36.
(27) Cook C, Heath F, Thompson RL. A meta-analysis of response rates in web- or internet-based surveys. Educational and Psychological Measurement 2000;60:821-36.
(28) Rosenbaum L. Conflicts of interest: part 1: Reconnecting the dots--reinterpreting industry-physician relations. N Engl J Med 2015;372:1860-4.

Chapter 8

Machiavellianism is associated with research misbehavior in Dutch biomedical scientists

Joeri K.Tijdink, Lex Bouter, Coosje L. S. Veldkamp, Peter M. van de Ven, Jelte M. Wicherts, Yvo M. Smulders

8

Under review 2016

Abstract

Relevance: Personality influences decision making and ethical considerations. Its influence on the occurrence of research misbehavior has never been studied. A large, quantitative study is important to determine whether personality can be identified as a risk factor for research misbehavior.

Objectives: This study aims to determine the association between personality traits and self-reported research misbehavior (fraud and questionable research practices). We hypothesized that narcissistic, Machiavellianistic and psychopathic traits are associated with research misbehavior.

Design: Cross-sectional study.

Setting: Biomedical scientists across hierarchical layers of 4 university medical centers in the Netherlands.

Participants: 535 Dutch biomedical scientists (response rate 65%), actively involved in scientific publishing.

Main outcome measures: We used validated personality questionnaires such as the Dark Triad (narcissism, psychopathy and Machiavellianism), Rosenberg's Self-Esteem Scale, the Publication Pressure Questionnaire (PPQ) and also demographic and job-specific characteristics to investigate the association with a composite self-reported research misbehavior severity score.

Results: Machiavellianism was positively associated (beta 1.28, CI 1.06 – 1.53) with self-reported research misbehavior while narcissism, psychopathy and self-esteem were not. Exploratory analysis revealed that, among persons in higher academic ranks (i.e. professors), narcissism and research misconduct were more severe ($p<0.01$ and $p<0.001$, respectively), and self-esteem scores and publication pressure were lower ($p<0.001$ and $p<0.01$, respectively) compared to PhD students.

Conclusions and Relevance: Machiavellianism may be a risk factor for research misbehavior. Narcissism and research misbehavior were more prevalent among biomedical scientists in higher academic positions. These results suggest that personality will have impact on research behavior and should be accounted for in fostering responsible conduct of research.

Introduction

Little is known about the psychology and personality of biomedical scientists. We like to think that scientists are open, eager to collaborate, self-confident, curious and creative. (1) However, there's anecdotal evidence that this is not universally so. (2) Success in science requires publishing in high Impact Factor journals and acquiring research grants, all in a hypercompetitive climate. This may seduce scientists to rush into print, cut corners, exaggerate findings and overstate the importance of their research (3). In addition, the so-called Dark Triad of personality, referring to Machiavellianism, Psychopathy, and Narcissism has been found to predict behaviors like abusive supervision and employee theft. (6) Self-esteem appears to be negatively associated with cheating, at least among students. (7;8)

The influence of personality traits on scientific practice is understudied. It has been shown that narcissism and cynicism show consistent negative relationships with aspects of ethical decision-making in future scientists (9) and these traits have been found to predict cheating in scholastic,(10) financial, (11), and work (12) settings. It is hence conceivable that these traits are associated with an increased likelihood to engage in research misbehavior (fraud and Questionable Research Practices or QRPs). Research misbehavior has received substantial attention in the last decade. (13) There is increasing evidence that research misbehaviors, specifically QRPs, are relatively common, seriously impact the scientific process, and compromises the validity of scientific results. (14)

Preliminary data indeed suggest that specific personality characteristics are indeed associated with scientific misbehavior. Psychiatrist Kornfeld (15) divided fraudulent scientists into different categories, suggesting that certain personality profiles are more common among fraudsters, and found a relationship between publication pressure and research misbehavior. This qualitative evidence was supported by quantitative results. (16)

In this study we aim to provide more insight in the psychology of research misbehavior. We hypothesize that some scientists are more vulnerable than others to commit research misbehavior. Specifically, we postulate that high self-esteem, Machiavellianism (a person's tendency to be unemotional, detached from conventional morality and hence inclined to deceive and manipulate others, to focus on unmitigated achievement and to give high priority to their own performances),(17) narcissistic traits (a person's tendency to pursue gratification from vanity or egotistic admiration, and to obtain recognition of their own attributes), (18) and psychopathic traits (enduring antisocial behavior, diminished empathy and remorse, and disinhibited or bold behavior) (19) are associated with research misbehavior. Furthermore, we will determine whether publication pressure and academic position influence the relation between personality traits and research misbehavior.

Materials and Methods

Participant selection and procedure

All 1833 biomedical scientists working in 4 university medical centers in the Netherlands were invited by email to participate in the survey. Scientists were eligible if they were sufficiently proficient in English, were scientifically active, and gave provided consent by following the link to the online questionnaire.

The research councils of the participating institutions provided e-mail addresses of all active scientists of a total of 9 departments. Two preclinical, three clinical (internal medicine, surgery and psychiatry) and two supportive departments (i.e. methodology, statistics) were selected. To create heterogeneity among the participants we also included the most and least publishing department per fte. The invitation e-mail (See digital supplement S1) explained the objective of the study, using neutral terms such as achievement, motivation, personality and scientific success, and provided them with a link to an anonymous online questionnaire (Digital supplement S2) on a protected website. The e-mail also included the name and a link to the e-mail address of the lead investigator (JT) to opt out for participation.

Scientists who did not respond within 2 weeks were sent two reminders. After the second reminder we asked invited participants who still did not respond to fill out a 15 second ultra-brief questionnaire to determine their reason for declining participation.

Survey characteristics and outcomes

The questionnaire contained, apart from demographic data, seven validated questionnaires (see digital Supplement S2). We used the Dark Triad (20) to measure Machiavellianism, narcissism and psychopathy. We used the Rosenberg Self-Esteem scale (21) to measure self-esteem. For the primary outcome we constructed a research misbehavior (fraud and QRPs) severity questionnaire yielding a composite research misbehavior severity score. This questionnaire was based on questionnaires used by other investigators, with additional items gathered from different landmark publications on research misbehavior. (14;16;22) It consisted of 22 different types of research misbehavior (see table 2). Survey respondents were asked to report to what extent they had committed specified types of research misbehavior during the past three years. Answers were given on a 5-point scale (never, once, occasionally, frequently, often). To construct a Research Misbehavior Severity Score (RMSS), the scores of the items were dichotomously translated (behavior yes/no) and items were assigned different weights. To construct this score, positive answers (committing the behavior at least 'once' in the past 3 years) to the most severe misbehavior questions (items 1, 2, 8, 9, 12, 15 and 19) were assigned three points, positive answers of the severe research misbehavior questions were assigned two points (items 4, 7, 10, 14, 16, 18 and 20) and positive answers of the moderate research misbehavior questions were assigned one point (items 3, 5, 6, 11, 13, 17, 21 and 22). Scores were added up to calculate the composite research misbehavior severity score (RMSS) (maximum range: 0-43).

In figure 1 we show the predefined analysis model. We postulate that 4 traits (narcissism, Machiavellianism, psychopathy and self-esteem) are related to the outcome variable (RMSS). We hypothesized that this relation could be modified or mediated by publication pressure or academic position. To measure publication pressure we used the validated publication pressure questionnaire (23). Academic position was operationalized by self-reported rank (PhD-students, postdoctorals, assistant professors, associate professors, and full professors).

Respondents provided demographic information on gender, age, marital status, number of children living at home, academic position, type of specialty, number of years working as a scientist, main professional activity (research, education, patient care or management).

Statistical Analysis

Associations between personality traits and the outcome measure (RMSS) were first tested by linear regression analyses using a separate regression model for each personality trait. In order to satisfy the normality assumption for the residuals the outcome variable (RMSS) was transformed and log(1+RMSS) was used as the dependent variable. Personality traits were standardized by subtracting their sample mean and dividing by the standard deviation. The exponentiated regression coefficients refer to the relative increase of geometric means of RMSS+1 associated with an increase of 1 standard deviation on the personality trait scale. Effect modification was assessed by including the candidate modifier (publication pressure or academic position) in the regression model together with its interaction with the personality trait at issue. A variable was considered to be an effect modifier if it showed a significant interaction ($p<0.05$) with the personality trait. To assess whether publication pressure and academic position mediated the relationship between personality traits and scientific misbehavior we divided the total effect into a direct effect of the personality trait and an indirect effect via the candidate mediator. A variable was considered a mediator when both the total effect and the indirect effect were significant. In the mediation analyses for academic position we used a probit-link to relate academic position to the personality trait.

In a secondary analysis, ANOVA tests were used to compare mean scores for personality traits, self-reported research misbehaviour and publication pressure

Figure 1. Predefined analysis model

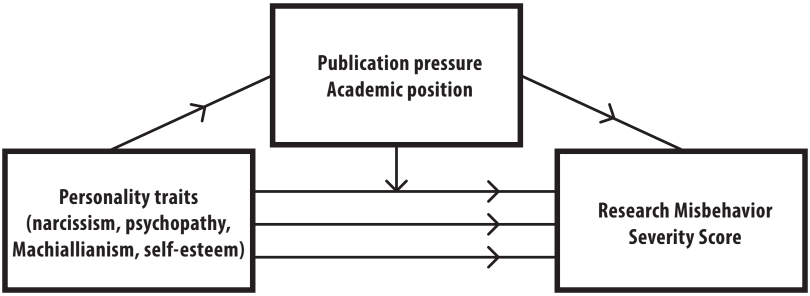


Table 1. Demographics

		N=535	%
Gender	Male	229	42.8
	Female	306	57.2
Age	<40	396	74
	>40	139	26
Academic Position	PhD student	303	56.6
	Postdoc, Assistant or Associate professors	177	33.1
	Full Professor	55	10.3
Years working as a scientist	0-4	220	41.1
	5-10	158	29.5
	11-15	46	8.6
	16-20	35	6.7
	21-25	26	4.7
	>25	49	9.2

Table 2. Items of the RMSS *Moderate misbehavior, ** severe misbehavior, *** most severe misbehavior

Research Misbehavior item	0 times (%)	Once (%)	Several times (%)	Regularly (%)	Always (%)
1. Modified the results or conclusions of a study under pressure from an organization that (co-)funded the research?***	520 (97.2)	9 (1.7)	6 (1.1)	0	0
2. To confirm a hypothesis, selectively deleted or changing data after performing data analysis? ***	510 (95.3)	18 (3.4)	7 (1.3)	0	0
3. Deleted data before performing data analysis? *	473 (88.4)	24 (4.5)	32 (6.0)	4 (0.7)	2 (0.4)
4. Concealed results that contradicted previous research you published?**	510 (95.3)	20 (3.7)	5 (0.9)	0	0
5. Used phrases or ideas of others without their permission?*	466 (87.1)	38 (7.1)	27 (5.0)	4 (0.7)	0
6. Used phrases or ideas of others without citation? *	464 (86.7)	33 (6.2)	32 (6.0)	5 (0.9)	1 (0.2)
7. Turned a blind eye to colleagues' use of flawed data or questionable interpretation of data?**	420 (78.5)	61 (11.4)	48 (9.0)	4 (0.7)	2 (0.4)
8. Fabricated data?***	533 (99.6)	1 (0.2)	0	0	1 (0.2)
9. Not published (part of) the results of a study?***	446 (83.4)	49 (9.2)	36 (6.7)	4 (0.7)	0
10. Deliberately not mentioned an organization that funded your research in the publication of your study?**	531 (99.3)	0	4 (0.7)	0	0
11. Added one or more authors to a report who did not qualify for authorship (honorary author)?*	213 (39.8)	130 (24.3)	150(28)	39 (7.3)	3 (0.6)
12. Selectively modified data after performing data analysis to confirm a hypothesis?***	514 (96.1)	16 (3.0)	5 (0.9)	0	0
13. Reported a downwardly rounded p value (e.g. reporting that a p value of .054 is less than .05)?*	524 (97.9)	7 (1.3)	3 (0.6)	1 (0.2)	0
14. Reported an unexpected finding as having been hypothesized from the start? **	429 (80.2)	63 (11.8)	39 (7.3)	4 (0.7)	0
15. Decided whether to exclude data after looking at the impact of doing so on the results?***	443 (82.8)	54 (10.1)	37 (6.9)	1 (0.2)	0
16. Decided to collect more data after seeing that the results were almost statistically significant?**	387 (72.3)	69 (12.9)	66 (12.3)	11 (2.1)	2 (0.4)
17. Omitted a contributor who deserved authorship from the author's list?*	521 (97.4)	7 (1.3)	6 (1.1)	1 (0.2)	0
18. Stopped collecting data earlier than planned because the result at hand already reached statistical significance without formal stopping rules?**	511 (95.5)	15 (2.8)	5 (0.9)	3 (0.6)	1 (0.2)
19. Deliberately failed to mention important aspects of the study in the paper?***	516 (96.4)	14 (2.6)	4 (0.7)	1 (0.2)	0
20. Not disclosed a relevant financial or intellectual conflict of interest?**	527 (98.5)	5 (0.9)	2 (0.4)	1(0.2)	0
21. Spread results over more papers than needed to publish more papers ('salami slicing')?*	440 (82.2)	53 (9.9)	29 (5.4)	13 (2.4)	0
22. Used confidential reviewer information for own research or publications?*	516 (96.4)	15 (2.8)	3 (0.6)	1 (0.2)	0

between respondents with different academic position. In case of a significant overall difference, pairs of groups were compared post hoc analyses using a Bonferroni correction.
Regression analyses and ANOVA tests were performed in SPSS version 22. Analyses of mediating variables were performed in M-Plus version 7 using the product of coefficients methods.

Results

In total, we used 1833 email addresses. Of these, 182 bounced because the email addresses no longer existed or were inactive. Of the remaining 1651, 1098 invitees opened the email. Among them, 715 started the survey (response rate 65%) and 537 completed the survey (completion rate 49%). We excluded 2 participants as they declared they were not scientifically active. The demographic data of the complete responders are summarized in table 1.

Research Misbehavior Severity Score (RMSS)

The items of the research misbehavior questionnaire are tabulated in table 2. Almost 60% of participants admitted they have added at least once authors without significant contribution. Fabrication and falsification were less common, although almost 5% admitted they had at least once selectively deleted data to confirm a hypothesis. The median RMSS score in the sample was 3 (range 0-39).

Machiavellianism predicts research misbehavior

Tables 3 and 4 present relations of personality traits with the research misbehavior severity score. Higher scores on Machiavellianism were significantly associated with higher scores of the research misbehavior severity score. There was a trend ($0.05 \leq p \leq 0.10$) for Narcissism and psychopathy having a similar association with research misbehavior. In multivariate analysis including all three personality traits, only Machiavellianism was significantly associated with the RMSS (data not shown).

Publication pressure as measured with the PPQ and academic position failed to significantly modify the relationship between the personality traits and research misbehavior severity: neither of them showed significant an interaction with any of the personality traits (see supplementary table T1).

Role of publication pressure and academic position

We considered whether publication pressure or academic position mediated the relation between the dark triad and RMSS scores. The results of the mediation analyses are displayed in table 3. In all mediation analyses PPQ and higher academic position were both found to be positively associated with RMSS scores (all $p < 0.001$).

Personality traits and demographic factors

Table 4 provides the means of the measured personality traits, PPQ and primary outcome (RMSS) stratified for academic position. Mean scores on narcissism and self-esteem were different between academic positions. Post-hoc tests revealed that both postdocs, assistant and associate professors, and professors had higher scores on narcissism (Bonferroni corrected $p < 0.05$ and $p < 0.01$, respectively) and lower scores on self-esteem compared to PhD-students (Bonferroni corrected $p < 0.05$ and $p < 0.01$, respectively). No significant differences in personality trait scores were found between postdocs and professors. Furthermore, the RMSS scores differed between the academic position groups and in post-hoc tests RMSS scores were found to be higher in professors and postdocs compared to PhD-students (Bonferroni corrected $p < 0.01$ and $p < 0.001$, respectively). No significant difference in RMSS was found between postdocs and professors. Finally, the publication pressure questionnaire score was found to differ between respondents with different academic positions. Post-hoc tests revealed a significant difference in PPQ scores between PhD-students and full professors (Bonferroni corrected $p < 0.05$) with professors reporting lower publication pressure.

Discussion

To our knowledge, this is the first study to investigate personality traits in relation to research misbehavior among biomedical scientists. Our results suggest that Machiavellianism is associated with self-reported research misbehavior. Secondary analyses reveal that narcissism, psychopathy and research misbehavior are positively associated with academic rank. Furthermore, higher

Table 3. Mediation analyses. Exponentiated regression coefficients, 95% confidence intervals and p-values for total, direct and indirect effects associated with 1 standard deviation increase in the personality trait.

	Mediation analysis				
		PPQ		Academic position	
	Total effect	Indirect	Direct	Indirect	Direct
Narcissism	1.08 (CI 1.00 - 1.16) p = 0.06	1.00 (CI 0.99 - 1.02) p = 0.76	1.07 (CI 1.00 - 1.16) p = 0.06	1.04 (CI 1.01 - 1.06) p = 0.005	1.04 (CI 0.96 - 1.12) p = 0.33
Psychopathy	1.08 (1.00, 1.16) p = 0.05 (ns)	1.00 (CI 0.98 - 1.01) p = 0.70	1.08 (CI 1.00 - 1.17) p = 0.04	1.01 (CI 0.99 - 1.03) p = 0.27	1.07 (CI 0.99 - 1.15) p = 0.08
Machiavellianism	1.12 (CI 1.04 - 1.21) p = 0.003	1.01 (CI 1.00 - 1.03) p = 0.07	1.11 (CI 1.03 - 1.20) p = 0.007	0.99 (CI 0.98 - 1.01) p = 0.90	1.12 (CI 1.04 - 1.21) p = 0.004
Self esteem	0.98 (CI 0.91 - 1.06) p = 0.60	1.01 (CI 0,99 - 1.02) p = 0.23	0.98 (CI 0.90 - 1.05) p = 0.46	0.96 (CI 0.94 - 0.99) p = 0.004	1.02 (CI 0.94 - 1.11) p = 0.63

8

Table 4. Academic position and personality traits. Table 4 shows the mean values of the dark triad: (Machiavellianism, narcissism, and psychopathy), self-esteem and PPQ sum score and comparison of means between groups using an ANOVA. Kruskal-Wallis test used to compare RMSS sum scores between groups.

		PhD-students (n=303)	Postdoctorals, assistant & associate professors (n=177)	Full professors (n=55)	ANOVA p-value
		Determinants			
Self-esteem	18.4 (CI 18.0 – 18.7)	18.8 (CI 18.4 – 19.3)	17.9 (CI 17.4 – 18.5)	16.9 (CI 16.1 – 17.8)	0.001
Narcissism	25.2 (CI 24.9 – 25.6)	24.7 ((CI 24.2 – 25.2)	25.4 (CI 25.1 – 26.3)	26.5 (CI 25.5 – 27.5)	0.002
Machiavellianism	25.0 (CI 24.6 – 25.3)	24.8 (CI 24.4 – 25.3)	25.4 (CI 24.8 – 26.0)	24.0 (CI 22.3 – 25.1)	0.09
Psychopathy	18.2 (CI 17.8 – 18.5)	18.0 (CI 17.6 – 18.5)	18.3 (CI 17.7 – 18.9)	18.7 (CI 17.7 – 19.8)	0.46
			Primary outcome measure		
RMSS		3.6 (CI 3.1 – 4.1) 2 (IQR: 1-5)	4.9 (CI 4.1 – 5.7) 4 (IQR: 1-7)	6.4 (CI 4.8 – 8.0) 5 (IQR: 2-9)	<0.001
			Candidate effect modifier		
PPQ sum score	42.2 (CI 42.0 – 43.1)	43.1 (CI 42.4 – 43.8)	42.2 (CI 41.3 – 43.3)	40.4 (CI 38.5 – 42.3)	0.017

academic positions are associated with both lower publication pressure and lower self-esteem as opposed to lower academic positions.

Although evidence from earlier research is lacking, our results seem largely in agreement with the qualitative narrative analysis of Kornfeld, (15) who gathered case histories of 146 individuals found guilty of research misbehavior and categorized them into different psychological traits. According to Kornfeld, scientific fraud is the product of a combination of individual personality traits and an intense fear of failure or the lure of academic and/or financial rewards. (15)Furthermore, most subjects declared an intense pressure to publish as the main reason for their behavior, reasoning that publications boost their career potential and financial rewards. Our study is larger and non-selective, and addresses a pre-defined hypothesis with quantitative data. In addition, Kornfeld's study had no comparison group of non-fraudulent scientists. It may also be that personality traits in fraudsters who were caught differ from those who are not (yet) caught.

In our analysis, Machiavellianism was associated with self-reported research misbehavior. Machiavellianism is best described as 'a person's tendency to be unemotional, detached from conventional morality and hence to deceive and manipulate others, focused on unmitigated achievement and high priority of own performances'. (17) This description intuitively explains that Machiavellianistic scientists more easily engage in research misbehavior. Moreover, the intellectual legacy of Niccolo Machiavelli confirms our findings as well. This is best illustrated by some of his quotes: 'Whosoever desires constant success must change his conduct with the times' (24) and 'One who deceives will always find those who allow themselves to be deceived' (25). Comparison with existing literature in the general population revealed that the levels of the subscales of the Dark Triad (including Machiavellianism) found in our study are comparable with the most recent literature, (20) suggesting that Dark Triad traits are not higher in biomedical researchers than in the general population. Our secondary analyses suggest that narcissistic and psychopathic traits are more common in higher academic ranks and that higher academics have less self-esteem. This may imply that the personality traits narcissism and psychopathy offer some kind of 'survival benefit' in academia. Publication pressure was also lower in higher academic ranks, which corresponds to earlier results. (16)

To our knowledge, this is the first study that investigates personality traits among biomedical researchers and their relation with major and minor research misbehaviors. The response rate was high compared to other online surveys. (26) We chose anonymous questionnaires because the questions on research misbehavior could cause hesitation to honest answering. (27;28)

The results of our study should, however, be interpreted with caution. First, internet-based questionnaires can be influenced by response bias, e.g. by attracting participants who are not engaged in research misbehavior and reticence by participants who have engaged in this type of behavior. To minimize this, we did not convey the purpose of this study in our invitational email and formulated the study goal in neutral terms (i.e. 'We invite you to participate in this brief questionnaire that addresses personal characteristics of biomedical scientists in relation to science practice', see Digital Supplement. S1). This, together with the high response rate suggest that the findings may be generalizable to the total population of biomedical researchers in the Netherlands (and possibly the rest of the industrialized world).

The research misbehavior severity score should also be interpreted with caution. The 22 items are all self-reported, are prone to different interpretation and were measured at one time point only. The composite score should also be cautiously interpreted as we composed the score, according to a self-designed (and thus arbitrary) one-dimensional score sheet and according to the input of earlier studies.

Taken together, although narcissism and psychopathy may be associated with research misconduct at first sight, the results suggest that of the examined traits, Machiavellianism is the personality trait that is most strongly associated with research misbehavior. These results may inform those involved in recruitment of scientific personnel, as well as people involved in scientific quality and integrity monitoring and those responsible for institutional research integrity policy and for responsible conduct of research training.

Supplementary

The Publication pressure questionnaire (PPQ)

This is a questionnaire analysing publication pressure. Please fill in to what extend you agree on the next statements

Likert Scale: 0 = totally disagree; 1 = disagree; 2 = no opinion; 3 = agree; 4 = totally agree

1. Without publication pressure, my scientific output would be of higher quality
2. My scientific publications contribute to better (future) medical care(R)
3. I experience my colleagues' assessment of me on the basis of my publications as stressful
4. I experience the publication criteria formulated by my university for my appointment or re-appointment as professor as a stimulus(R)
5. Publication pressure puts pressure on relationships with fellow researchers
6. I suspect that publication pressure leads some colleagues (whether intentionally or not) to color data
7. The validity of medical world literature is increased by the publication pressure in scientific centers(R)
8. Publication pressure leads to serious worldwide doubts about the validity of research results
9. In my opinion, the pressure to publish scientific articles has become too high
10. The competitive scientific climate stimulates me to publish more(R)
11. My colleagues judge me mainly on the basis of my publications
12. Fellow scientists maintain their clinical and teaching skills well, despite publication pressure(R)
13. I cannot confide innovative research proposals to my colleagues
14. Publication pressure harms science

(R) reversed questions

The Rosenberg Self Esteem Scale

Below is a list of statements dealing with your general feelings about yourself. If you strongly agree, circle **SA**. If you agree with the statement, circle **A**. If you disagree, circle **D**. If you strongly disagree, circle **SD**.

SA= strongly agree, A= agree, D= disagree, SD= strongly disagree

1. In general I am satisfied with myself
2. Sometimes I think I am good for nothing (R)
3. I own several good qualities
4. I am capable of doing all sorts of things just as well as most other people
5. In my opinion I have not much to be proud of (R)
6. Sometimes I feel really useless (R)
7. I find myself just as much worth as others
8. I wish I had a little more selfrespect (R)
9. All things considerated I tend to call myself a loser (R)
10. I am fairly pleased with myself

(R) Reversed Items

The Dark Triad

The next questionnaire consists of twenty seven guidelines. Please rate your agreement or disagreement with each item using the following guidelines: If you strongly agree, circle **SA**. If you agree with the statement, circle **A**. If you neither agree nor disagree circle if you disagree with the statement, circle **D**. If you strongly disagree, circle **SD**.

1= strongly agree, 2= agree, 3=neutral 4= disagree, 5= strongly disagree

1. It's not wise to tell your secrets.
2. People see me as a natural leader.
3. I like to get revenge on authorities.
4. Generally speaking, people won't work hard unless they have to
5. I hate being the center of attention. R
6. I avoid dangerous situations. R

7. Whatever it takes, you must get the important people on your side.
8. Many group activities tend to be dull without me.
9. Payback needs to be quick and nasty.
10. Avoid direct conflict with others because they may be useful in the future.
11. I know that I am special because everyone keeps telling me so.
12. People often say I'm out of control.
13. It's wise to keep track of information that you can use against people later.
14. I like to get acquainted with important people.
15. It's true that I can be mean to others.
16. You should wait for the right time to get back at people.
17. I feel embarrassed if someone compliments me. R
18. People who mess with me always regret it.
19. There are things you should hide from other people because they don't need to know.
20. I have been compared to famous people.
21. I have never gotten into trouble with the law. R
22. Make sure your plans benefit you, not others.
23. I am an average person. R
24. I like to pick on losers
25. Most people can be manipulated.
26. I insist on getting the respect I deserve.
27. I'll say anything to get what I want.

R = reversed item

Research Misbehavior Severity Score

In your work as a scientist, have you shown, even if it has been only on a sigle occasion, any of the following behaviors in the last three years?

Table T1. Exponentiated regression coefficients Exp(beta) for linear regression of RMSS on personality traits and results of tests effect modification by PPQ and academic position (to clarify the beta scores: an increase of 1 standard deviation in Machiavellianism is associated with an increase of 12% in the geometric mean of RMSS+1)

1.	Modified the results or conclusions of a study under pressure from an organization that (co)funded the research?
2.	To confirm a hypothesis, selectively deleted or changing data after performing data analysis?
3.	Deleted data before performing data analysis?
4.	Concealed results that contradicted previous research you published?
5.	Used phrases or ideas of others without their permission?
6.	Used phrases or ideas of others without citation?
7.	Turned a blind eye to colleagues' use of flawed data or questionable interpretation of data?
8.	Fabricated data?
9.	Not published (important part of) the results of a study?
10.	Deliberately not mentioned an organization that funded your research in the publication of your study?
11.	Added one or more authors to a report who did not qualify for authorship (honorary author)?
12.	Selectively modified data after performing data analysis to confirm a hypothesis?
13.	Reported a downwardly rounded p value (e.g. reporting that a p value of .054 is less than .05)?
14.	Reported an unexpected finding as having been hypothesized from the start?
15.	Decided whether to exclude data after looking at the impact of doing so on the results?
16.	Decided to collect more data after seeing that the results were almost statistically significant?
17.	Omitted a contributor who deserved authorship from the author's list?
18.	Stopped collecting data earlier than planned because the result at hand already reached statistical significance without formal stopping rules?
19.	Deliberately failed to mention important aspects of the study in the paper?
20.	Not disclosed a relevant financial or intellectual conflict of interest?
21.	Spread results over more papers than needed to publish more papers ('salami slicing')?
22.	Used confidential reviewer information for own research or publications?

8

Supplementary Tables

		Effect modification	
		PPQ	Academic position
	Exp(beta) 95% CI p-value	**p-value interaction**	**p-value interaction**
Narcissism	1.08 (CI 1.00 - 1.16) p = 0.06	0.70	0.07
Psychopathy	1.08 (CI 1.00 - 1.16) p = 0.05	0.44	0.68
Machiavellianism	1.12 (CI 1.04 - 1.21) p = 0.003	0.47	0.95
Self esteem	0.98 (CI 0.91 - 1.06) p = 0.60	0.20	0.20

Reference List

(1) Isohanni M, Kotiranta H. What is a top scientist like? Nord J Psychiatry 2006;60:189.

(2) Fetzer J. Wolfish behavior...the successful scientist as leader of a pack or a lone wolf. Anal Bioanal Chem 2003;377:785-786.

(3) Alberts B, Kirschner MW, Tilghman S, Varmus H. Rescuing US biomedical research from its systemic flaws. Proc Natl Acad Sci U S A 2014;111:5773-5777.

(4) Feist G.J. Psychology of Science as a New Subdiscipline in Psychology. Current Directions in Psychological Science 2011;20:330-334.

(5) Gino F, Ariely D. The dark side of creativity: original thinkers can be more dishonest. J Pers Soc Psychol 2012;102:445-459.

(6) O'Boyle EH, Forsyth DR, Banks GC, McDaniel MA. A meta-analysis of the Dark Triad and work behavior: a social exchange perspective. J Appl Psychol 2012;97:557-579.

(7) Aronson E, Mettee DR. Dishonest behavior as a function of differential levels of induced self-esteem. J Pers Soc Psychol 1968;9:121-127.

(8) Ward D.A. Self-Esteem and Dishonest Behavior Revisited. The Journal of Social Psychology 1986;126:709-713.

(9) Antes A.L., Brown R.P., Murphy S.T. et al. Personality and Ethical Decision-Making in Research: The Role of Perceptions of Self and others. Journal of Empirical Research on Human Research Ethics 2007;2:15-34.

(10) Nathanson C., Paulhus D.L., Williams K.M. Predictors of a behavioral measure of scholastic cheating: Personality and competence but not demographics. Contemporary Educational Psychology 2006;31:97-122.

(11) Jones D.N. Risk in the face of retribution: Psychopathic individuals persist in financial misbehavior among the Dark Triad. Personality and individual Differences 2014;67: 109-113.

(12) Cohen TR, Panter AT, Turan N, Morse L, Kim Y. Moral character in the workplace. J Pers Soc Psychol 2014;107:943-963.

(13) Wicherts JM. Psychology must learn a lesson from fraud case. Nature 2011;480:7.

(14) Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PLoS One 2009;4:e5738.

(15) Kornfeld DS. Perspective: research misconduct: the search for a remedy. Acad Med 2012;87:877-882.

(16) Tijdink JK, Verbeke R, Smulders YM. Publication pressure and scientific misconduct in medical scientists. Journal of Empirical Research and Human Research Ethics 2014;9:64-71.

(17) Christie R., Geis F. Studies in Machiavellianism. NY: Academic Press, 1970.

(18) Clarkin J.F., Fonagy P., Gabbard G.O. Narcissistic personality disorder. Psychodynamic Psychotherapy for Personality Disorders: A clinical Handbook. 1st ed. Arlington: American Psychiatric Publishing; 2010; p. 263.
(19) Skeem JL, Polaschek DL, Patrick CJ, Lilienfeld SO. Psychopathic Personality: Bridging the Gap Between Scientific Evidence and Public Policy. Psychological Science in the Public Interest 2011;12:95-162.
(20) Jones DN, Paulhus DL. Introducing the short Dark Triad (SD3): a brief measure of dark personality traits. Assessment 2014;21:28-41.
(21) Rosenberg M. Society and the adolescent self-image. Princeton, NJ: Princeton Press, 1965.
(22) Martinson BC, Anderson MS, DeVries R. Scientists behaving badly. Nature 2005;435:737-738.
(23) Tijdink JK, Smulders YM, Vergouwen AC, de Vet HC, Knol DL. The assessment of publication pressure in medical science; validity and reliability of a Publication Pressure Questionnaire (PPQ). Qual Life Res 2014;23:2055-2062.
(24) Detmold C.E. The Historial, political, and diplomatic writings of Niccolo Machiavelli. Boston: J.R. Osgood and company, 1882.
(25) Niccolo Machiavelli. The Prince. 1515;p.85.
(26) Cook C, Heath F, Thompson RL. A meta-analysis of response rates in web- or internet-based surveys. Educational and Psychological Measurement 2000;60:821-836.
(27) Joinson A. Social desirability, anonymity, and Internet-based questionnaires. Behavior Research Methods, Instruments, & Computers 1999;31:433-438.
(28) Tourangeau R, Smith T.W. Asking sensitive questions the impact of data collection mode, question format, and question context. Public opinion quarterly 1996;60: 275-304.

Chapter 9

Publiphilia Impactfactorius: a new psychiatric syndrome among biomedical scientists? [1]

Joeri K Tijdink, Yvo M Smulders, Lex M Bouter

9

1 The text was written for the BMJ 2015 Christmas issue and is meant as a 'tongue-in-cheek' article that is based on real data. Unfortunately, it was not accepted for publication after peer-review. It will be submitted to another journal.

Abstract

Objective: To explore clusters of personality traits among biomedical scientists, and to associate the clusters with academic position and research misbehaviour

Design: Cross-sectional study with cluster analysis of personality traits among biomedical scientists.

Setting: A stratified sample of Dutch biomedical scientists working in academic medical centres.

Participants: 537 active biomedical scientists completed a web-based survey (response rate 65%).

Main outcome measures: The NEO-BIG5, Rosenberg Self-esteem, Achievement Motivation Inventory and the Dark Triad (narcissistic, Machiavellianistic and psychopathic personality traits) were measured by validated questionnaires. Self-reported research misconduct was assessed via a separate questionnaire.

Results: Cluster analysis revealed the existence of three personality clusters among biomedical scientists: the 'perfectionist', the 'ideal son-in-law' and the 'sneaky grandiose'. The latter cluster showed a consistent set of (subclinical) personality traits such as narcissism, psychopathy and Machiavellianism, that are indicative of the presence of a mental disorder, but could not be classified as such in terms of the DSM-IV TR or ICD-10. Male gender, higher academic hierarchical position, perceived publication pressure and, importantly, self-reported scientific misbehaviour were associated with the 'sneaky grandiose' personality cluster.

Conclusions: These findings suggest that biomedical scientists in the 'sneaky grandiose' personality cluster have a relatively high propensity to engage in research misbehaviour. A small proportion of the 'sneaky grandiose' might suffer from a psychiatric condition characterized by pathological preoccupation with publishing and being cited. We therefore propose to name this syndrome 'Publiphilia Impactfactorius' (PI), and we suggest this affliction should be considered

in revised versions of DSM5 and ICD-10. We provide tentative diagnostic criteria for PI. Early identification and intensive treatment or, alternatively, expulsion and abandoning of colleagues who suffer from PI may prevent further accumulation of research waste.

Introduction

Personality traits differ significantly between professionals, and the biomedical field is no exception. (1;2) Biomedical scientists in particular have an increasingly doubtful reputation, mainly for producing a large amount of irrelevant and unreliable 'research waste' (3) with a view to enhancing their career perspectives and boosting their ego. However, whether specific clusters of personality traits are indeed typical for biomedical scientists is unknown. This can be of particular interest since personality traits impact on behavior, and sloppy science or even scientific misconduct may be linked to specific clusters of personality traits. Scientists' personality traits can thus potentially inform not only the selection of candidates for academic positions, but also targeted prevention programmes or even the decision to reject individuals.

In this study, we explore the personality traits of biomedical researchers and perform a cluster analysis to identify common combinations of such traits. We use validated personality questionnaires. Our secondary objective was to assess whether personality clusters, if they exist, are associated with personal and job-specific characteristics, research misbehavior and perceived publication pressure with a view to ultimately describe a new psychiatric syndrome.

Materials and Methods

Participant's selection and procedure

1833 biomedical scientists working in four medical university centres in the Netherlands were invited to participate in our web-based survey. Scientists were eligible to participate if they were able to read English, were scientifically active and gave informed consent by following the link to the online questionnaire.

E-mail addresses of the scientists were obtained via the research councils of the participating institutions. We collected e-mail addresses of scientists from nine departments per institution (two preclinical departments, three clinical departments (internal medicine, surgery and psychiatry), two supportive departments (ie epidemiology, public health), and the most and least publishing (per fte) department). We sent an invitation e-mail to explain the objective of the study, using neutral terms such as 'achievement', 'motivation', 'personality' and 'scientific success', and provided them with a link to an anonymous online questionnaire on a protected website. Scientists who did not respond within 2 weeks were sent 2 reminders. After the second reminder we asked invited participants who still did not respond to fill in an ultra-brief survey to determine their reason for declining participation.

Ethical approval

The ethical Review Committee of Vrije University medical centre (VUmc) approved the protocol and confirmed that the Medical Research Involving Human Subjects Act (WMO) does not apply. In the email to the participants, we explicitly stated that full protection of their identity was guaranteed.

Survey characteristics and outcomes

The survey contained, apart from demographic data and job specific questions, six (validated) questionnaires.

To measure personality traits, we used the Dark Triad (testing narcissism, Machiavellianism and psychopathy), (4) the NEO Big Five testing neuroticism, extraversion, openness, conscientiousness and agreeableness), (5) the Achievement Motivation Inventory (6) and the Rosenberg Self-Esteem questionnaire. (7)

Machiavellianism is often described as 'to be unemotional, detached from conventional morality and prone to deceive and manipulate others, focused on unmitigated achievement and high priority of own performances'. (8) Narcissism is referred to a 'tendency to pursue gratification from vanity or egotistic admiration and recognition of one's own attributes' (9) and psychopathy is characterized by 'enduring antisocial behavior, diminished empathy and remorse, and disinhibited or bold behavior'. (10)

Research misbehaviour was measured by a composite scale (See supplementary table S2) based on questionnaires used by investigators with additional items gathered from different landmark publications on research misbehaviour. (11-13) To measure publication pressure, we used the validated Publication Pressure Questionnaire (PPQ). (14)
Respondents provided demographic information on gender, age, academic position (PhD student, postdoc, (assistant, associate and full) professor), type of specialty; years working as a scientist, main professional activity (research, education, patient care or management), and Hirsch index. (15) The survey was primarily designed to relate personality traits with research misbehaviour (manuscript submitted for publication).

Statistical Analysis

We used cluster analysis to explore the existence of different personality clusters of biomedical scientists, performed with SPSS version 20. With this technique participants were clustered into groups that resemble each other more than they resemble the participants outside the group at issue. First we transformed the scores from the questionnaires into z-scores and then we fitted 2-cluster, 3-cluster and 4-cluster solutions, according to the standard methods described for cluster analysis. (16;17)

As a validation procedure, we conducted a split-half cross-validation for the total group of respondents. Cohen’s kappa was used to compare the agreement between estimated and predicted values of the three clusters.

Analysis of Variance was used to compare clusters. Associations were explored between the personality clusters and the demographic and job specific characteristics (including the research misbehaviour severity score, see supplementary table S1).

Results

In total, we used 1833 email addresses. Of these, 182 bounced because the address no longer existed or was inactive. Of the remaining 1651, 1098 invitees opened the email, 715 started the survey (response rate 65%) and 537 completed the survey (completion rate 49%). We excluded 2 participants who declared they were not scientifically active.

The demographic data of the complete responders are summarized in table 1. 177-55, 33.1% 10.3%

Cluster analysis

Cluster analyses revealed that a 3-cluster model derived from 6 personality questionnaires offered optimal discrimination (ANOVA p<0.001 for all scales). These questionnaires were: the Rosenberg Self-Esteem questionnaire, the neuroticism subscale of the Neo-Big 5, the three subscales of the Dark Triad (Machiavellianism, narcissism and psychopathy subscales) and the Achievement Motivation Inventory. Table 2 specifies the three clusters.

Cluster 1 (n=140) is characterized by high levels of neuroticism ('emotionally reactive and vulnerable to stress' (18)), self-esteem and achievement motivation. These scientists have relatively low levels of narcissism. We decided to label this cluster as 'Perfectionists'.

Cluster 2 (n=192) has relatively low scores on self-esteem and achievement motivation. They are honest, 'easy going', and have the lowest scores on neuroticism ('They tend to be calm, emotionally stable, and free from persistent negative feelings'). (19) These biomedical scientists also have the lowest scores in narcissism, Machiavellianism and psychopathy. By consensus of our mothers, we decided to call this cluster 'Ideal son-in-law'.

The third cluster (n=205) is characterized by the highest levels of Machiavellianism, narcissism and psychopathy and with low self-esteem scores and a relative low motivation to achieve. After hefty deliberation, we decided to label this cluster as the 'Sneaky grandiose'.

Table 1. Demographics

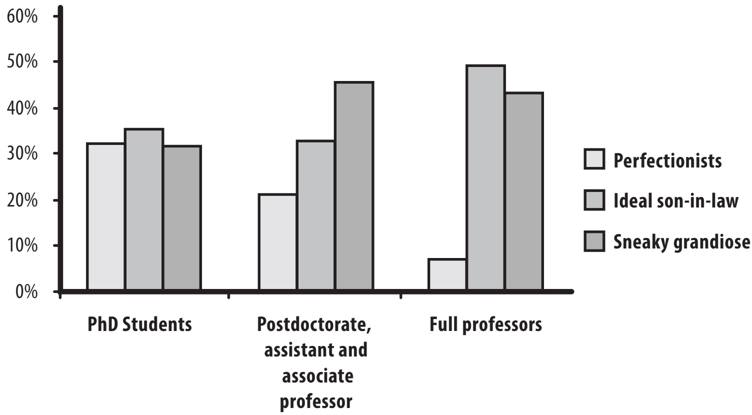
		N=535	%
Gender	Male	229	42.8%
	Female	306	57.2%
Age	<40	396	74%
	>40	139	26%
Academic Position	PhD student	303	56.6%
	Postdoc, associate and assistant professor	177	33.1%
	Full professors	55	10.3%
Years working as a scientist	0-4	220	41.1%
	5-10	158	29.5%
	11-15	46	8.6%
	16-20	35	6.7%
	21-25	26	4.7%
	>25	49	9.2%

9

Table 2. The mean scores, 95% CI and the z-scores for the 6 personality traits are provided for both the total group of participants and for the 3 personality clusters. According to cluster analysis methods, cluster differences with ANOVA were statistically significant for all 6 traits ($p<0.001$)

		Clusters of biomedical scientists					
	Mean (95% CI) (n=535)	I (n=140)		II (n=193)		III (n=202)	
		Mean (95% CI)	z-scores	Mean (95% CI)	z-scores	Mean (95% CI)	z-scores
Narcissism (Range 13-35)	25.2 (CI 24.9 – 25.6) IQR 22-27	22.5(CI 21.9 – 23.1) IQR 20-25	-0.66	23.8(CI 23.2 – 24.2) IQR 22-26	-0.36	28.5(CI 28.1 – 28.8) IQR 27-30	0.79
Machiavellianism (Range 9-38)	25.0 (CI 24.6 – 25.3) IQR 21-26	26.0(CI 25.3 – 26.6) IQR 23-27	0.24	21.7(CI 21.2 – 22.1) IQR 19-24	-0.78	27.4 (CI 26.9 – 27.9) IQR 25-30	0.58
Psychopathy (Range 9-30)	18.2 (CI 17.8 – 18.5) IQR 14-20	19.1(CI 18.4 – 19.7) IQR 17-21	0.21	15.2 (CI 14.8 – 15.6) IQR 13-17	-0.75	20.5(CI 20.0 – 20.9) IQR 18-23	0.56
Achievement Motivation (Range 14-35)	20.9 (CI 20.6 -21.3) IQR 17-23	23.5(CI 22.8 – 24.1) IQR 21-26	0.64	20.7(CI 20.2 – 21.2) IQR 18-24	-0.05	19.4 (CI 18.9 – 19.8) IQR 17-21	-0.39
Self Esteem (Range 10-32)	18.4 (CI 18.0 – 18.7) IQR 15-20	22.3(CI 21.8 – 22.8) IQR 20-24	1.02	17.1(CI 16.7 – 17.6) IQR 15-20	-0.32	16.8 (CI 16.4 – 17.2) IQR 15-19	-0.40
Neuroticism (Range 4-20)	10.1 (CI 9.9 – 10.4) IQR 7-13	12.5(CI 12.0 – 13.0) IQR 10-15	0.73	8.8 (CI 8.4 – 9.2) IQR 7-11	-0.41	9.8 (CI 9.3 – 10.2) IQR 8-12	-0.11
		Perfectionist		Ideal son-in-law		Sneaky grandiose	

Figure 1a. respondents (in %) per academic rank, stratified for the 3 clusters (p<0.001)



As a validation procedure, we conducted a split-half cross-validation of the data. This yielded Cohen kappas of 0.826 and 0.845 with p-values p< 0.0001), which can be interpreted as large.

Figure 1b. The H-index, stratified for the 3 clusters (p=NS).

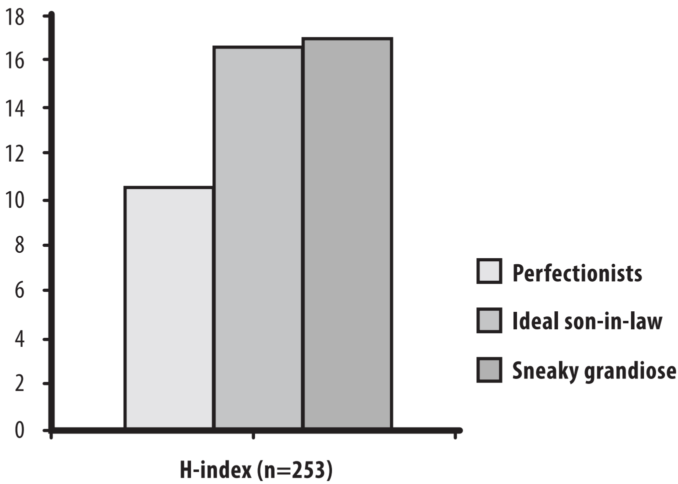


Figure 1c. Research misbehaviour severity score, stratified for the 3 clusters ($p<0.05$).

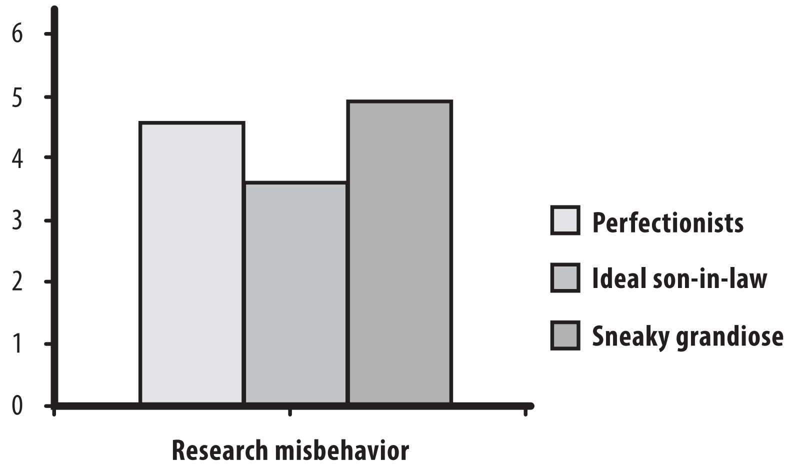
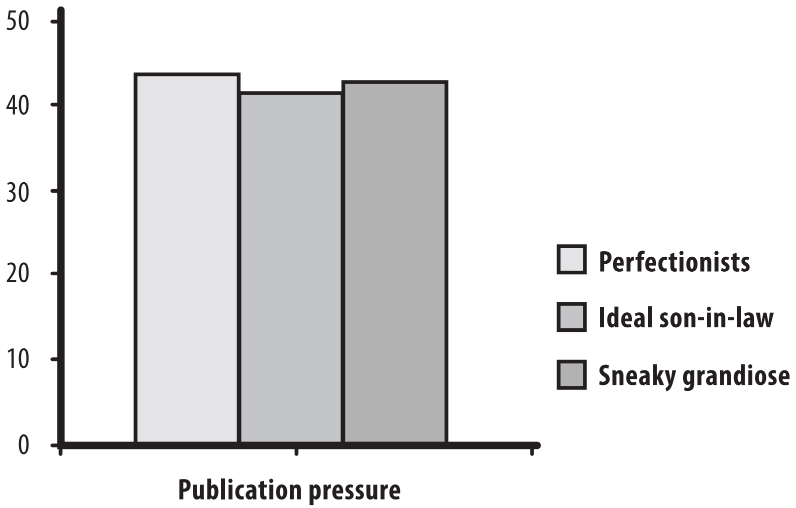


Figure 1d. Perceived publication pressure, stratified for the 3 clusters ($p<0.01$).



Who are these people?

Demographic and job-specific characteristics were compared between the clusters (see figure 1a-d, and supplementary table S2). 'Sneaky grandiose' was the dominant phenotype among men, whereas 'Ideal son-in-law' was most prevalent in female biomedical scientists. This suggests that 'Ideal daughter-in-law' may be a more appropriate label for this cluster. Perfectionists turned out to be often under 40 years of age or PhD students. Full professors were rare among the Perfectionists and common among the Sneaky grandiose. That may explain why Perfectionists had a relatively low H-index. Personality clusters were evenly distributed in scientists under 40 years of age/PhD student, but progressively few perfectionists were found among biomedical scientists who were more senior, in terms of age, academic position or Hirsch index.

Research misbehaviour severity and publication pressure scores differed between the clusters of biomedical scientists (see figure 1c and d). The ideal sons-in-law had lowest scores on the research misconduct severity score, whereas the sneaky grandiose cluster had the highest scores. The perfectionists reported the highest publication pressure.

Discussion

Salient findings and interpretation

To our knowledge, this is the first study that classifies personality traits in biomedical scientists. We identified three personality clusters. To make a career in biomedical science (more postdoctoral and professor positions, higher H-index), the data suggest, you need to be either an ideal son-in-law or a sneaky grandiose. The sneaky grandiose, however, display narcissistic, Machiavellianistic and psychopathic traits, have low self-esteem and high neuroticism. Their personality may predispose them to scientific misbehaviour. Hence we suggest targeting ideal sons (or daughters)-in-law for future key positions in biomedical science.

A previous study on personality traits of biomedical scientists included only those who were found guilty of research misconduct. (20) Hence, selected participants were all 'rotten apples', not representative of the whole spectrum of biomedical scientists. Nevertheless, some similarities are arresting. That study also unearthed (among other profiles) a personality profile called 'the grandiose', with a similar pattern of personality traits. No other studies to date are available on personality traits or personality clusters in biomedical or other scientists. Why are sneaky grandiose overrepresented in higher academic ranks? One possibility is that the character of many biomedical scientists evolves into this phenotype after prolonged exposure to a hostile environment, which includes perverse incentives, (21) hypercompetition, (22) and many bad examples and cheating role models. (13) The alternative explanation is that it is simply a matter of selection, where perfectionists are the first to be expelled from academia, and the sneaky grandiose have only ideal sons-in-law left to compete with. Without longitudinal studies, these questions are impossible to answer with certainty. However, narcissistic and psychopathic personality traits are predominantly genetically determined, (23) suggesting that evolution of such traits within a relatively short period of adulthood is unlikely to play a large role.

Comparison with existing literature in normal population revealed that the levels of the subscales of the Dark Triad (including Machiavellianism) are comparable with the most recent literature (4) suggesting that the traits are no higher in biomedical researchers than in the general public. However, the sneaky grandiose cluster has high levels of the three subscales of the Dark Triad compared with general public. (4)

One strength of our study is that the split-half cross-validation suggested high validity. Furthermore, we included a relatively large number of participants and had a high response rate (65%) compared to average response rates in web-based surveys. (24;25) Moreover, respondents were blinded to the primary objective of this study, which makes it implausible that response bias has influenced the results. Although we felt that it was highly unlikely that our study has limitations, some have to bear in mind. Some limitations have to be taken into account. This includes the cross sectional design of our study and the lack of a clear theoretical base for the findings. This makes longitudinal causal inferences regarding personality traits troublesome.

Another important bias is in the profile itself. Since the filthy grandiose profile is easily engaged in research misconduct, we should take that into account when we analyse and interpret the data; it will be very likely that respondents belonging in this cluster were not completely sincere while participating in our survey and might have made up their answers. After adjusting for this bias, it's likely that the real levels of narcissism and psychopathy will turn out to be much higher'

Unravelling a psychiatric disorder?

Our cluster analysis identifies a large proportion of medical scientists, mostly elderly males in high positions, belonging to the 'sneaky grandiose' personality cluster. It is certainly conceivable that a subset of the 'sneaky grandiose' have extreme levels of the Dark Triad traits as seen in the interquartile ranges within this cluster (see table 2). Some of the characteristics of sneaky grandiose scientists resemble features of the narcissistic and psychopathic personality disorders, which are existing classifications in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV TR) and the International Statistical Classification of Diseases (ICD-10). However, no psychiatric disorder describes the complete phenotype of extreme sneaky grandiose, whose key symptoms are described in Box 1. Since individuals suffering from more extreme features of the sneaky grandiose may pose a heavy burden to themselves and their surroundings, we propose a new psychiatric disorder: Publiphilia Impactfactorius (PI). We suggest this affliction should be considered while revising the DSM-IV TR or ICD-10. Auspiciously, the abbreviation of the syndrome parallels that of Principal Investigator.

How can you recognize Publiphilia Impactfactorius?

Individuals that suffer from PI can be recognized by obsessive preoccupation with citation indices, a strong urge to publish in high impact factor journals, profound despair and tantrum episodes after rejection of a manuscript, paranoid thoughts and envy towards colleagues, obsessive focus on authorship ranks, greed for higher academic positions, and a propensity to cut corners or worse. Some even have rage attacks after noticing that their Hirsch index had not risen since they looked at it a few days earlier. They are very much afraid of failure and believe that the end always justifies the means. They make tactical (for them beneficial) decisions in research collaborations, manipulate others to get things done and if needed, they are intentionally nasty and rude to coworkers, especially to junior colleagues. They easily lie and deceive to get ahead (see Box 1).

What can you do?

For the short term it is advisable to make significant changes to selection and promotion criteria. Researchers with PI might not be the talented new colleague of your preference that will give your department an honest boost in collaboration and trustworthiness. What institutions probably need – especially in leadership positions – are sincere, quiet, honest, trustworthy high achievers. Selection procedures should aim at recruiting ideal sons-in-law. To accomplish this, board members and head of research departments should bring their mother-in-law to the job interviews to select the ideal candidate. Or they can also directly appoint one of the female candidates, as PI seems to be very rare among them.

Diagnostic Criteria for Publiphilia Impactfactorius

A Five (or more) of the following symptoms are present:

1. Has a grandiose sense of self-importance (exaggerates scientific achievements and talents, demands admiration from inferiors)
2. Is preoccupied with fantasies of publications in high impact factor journals (New England Journal of Medicine and higher)
3. Selfish in all professional behavior, never altruistic
4. Lack of remorse, being indifferent to or rationalizing having hurt, mistreated or stolen from colleagues or co-authors
5. Manipulative in all professional relations and often using confidential information against colleagues
6. Can be impulsive, out of control and emotionally unstable, especially when coworkers contradict their beliefs and opinions
7. Continuously comparing their H-index with others with a view to confirm their own grandiosity. If colleagues have higher H-indexes, this can cause extreme envy and disgust
8. Emotional dependence on frequent publishing
9. Unable to cope with rejections of manuscripts or grant proposals

B. Evidence for having engaged in at least one of the following questionable research practices[1]

1. Turned a blind eye to colleagues' use of flawed data or questionable interpretation of data
2. Frequently demanding honorary and guest authorships without contribution
3. Decided to collect more data after seeing that the results were almost statistically significant
4. Not disclosed a relevant financial or intellectual conflict of interest

C. Having had at least 5 years of scientific work experience

D. Evidence that the pervasive pattern of symptoms was not displayed (yet existent) at the onset of his scientific career

E. Having a higher academic rank such as assistant, associate or full professor

[1] These 4 research misbehaviours were independently associated with higher incidences in the sneaky grandiose cluster ($p<0.05$).

9

Supplementary

Supplementary Table S1. The 22 items of the research misbehaviour severity score (RMSS) and their incidences in the study

Research Misbehaviour item	0 times (%)	Once (%)	Several times (%)	Regularly (%)	Always (%)
1. Modified the results or conclusions of a study under pressure from an organization that (co-)funded the research?	520 (97)	9 (2)	6 (1)	0	0
2. To confirm a hypothesis, selectively deleted or changing data after performing data analysis?	510 (95)	18 (3)	7 (1.3)	0	0
3. Deleted data before performing data analysis?	473 (88)	24 (4)	32 (6)	4 (0.7)	2 (0.4)
4. Concealed results that contradicted previous research you published?	510 (95)	20 (4)	5 (1)	0	0
5. Used phrases or ideas of others without their permission?	446 (87)	36 (7)	27 (5)	2 (0.7)	0
6. Used phrases or ideas of others without citation?	464 (86)	33 (6)	32 (6)	5 (0.9)	1 (0.2)
7. Turned a blind eye to colleagues' use of flawed data or questionable interpretation of data?	420 (78)	61 (11)	48 (9)	4 (0.7)	1 (0.4)
8. Fabricated data?	533 (99)	1 (0.2)	0	0	1 (0.2)
9. Not published (part of) the results of a study?	446 (83)	49 (9)	36 (7)	1 (0.7)	0
10. Deliberately not mentioned an organization that funded your research in the publication of your study?	531 (99)	0	4 (0.7)	0	0
11. Added one or more authors to a report who did not qualify for authorship (honorary author)?	213 (39)	130 (24)	150 (28)	39 (7)	1 (0.6)
12. Selectively modified data after performing data analysis to confirm a hypothesis?	514 (96)	16 (3)	5 (1)	0	0
13. Reported a downwardly rounded p value (e.g. reporting that a p value of .054 is less than .05)?	524 (98)	7 (1)	2 (0.6)	1 (0.2)	0
14. Reported an unexpected finding as having been hypothesized from the start?	429 (80)	63 (12)	39 (7)	1 (0.7)	0
15. Decided whether to exclude data after looking at the impact of doing so on the results?	443 (83)	54 (10)	37 (7)	1 (0.2)	0
16. Decided to collect more data after seeing that the results were almost statistically significant?	387 (72)	69 (13)	66 (12)	11 (2)	2 (0.4)
17. Omitted a contributor who deserved authorship from the author's list?	521 (97)	7 (1)	6 (1)	1 (0.2)	0
18. Stopped collecting data earlier than planned because the result at hand already reached statistical significance without formal stopping rules?	511 (95)	15 (3)	5 (1)	3 (0.6)	1 (0.2)
19. Deliberately failed to mention important aspects of the study in the paper?	516 (96)	14 (3)	4 (0.7)	1 (0.2)	0
20. Not disclosed a relevant financial or intellectual conflict of interest?	527 (98)	5 (1)	2 (0.4)	1 (0.2)	0
21. Spread results over more papers than needed to publish more papers ('salami slicing')?	440 (82)	53 (10)	29 (5)	13 (2)	0
22. Used confidential reviewer information for own research or publications?	516 (96)	15 (3)	3 (0.6)	1 (0.2)	0

Appendix to the table S1 clarifying the severity score:

The questionnaire consisted of 22 different types of research misbehaviour (see table S1). Survey respondents were asked to report to what extent they have been committed specified types of research misbehaviour during the past 3 years. Answers were given on a 5-point scale (never, once, occasionally, frequently, often). To construct a misbehaviour severity score of research misbehaviour, the items were assigned different points to calculate a composite research misbehaviour severity score (RMSS). To construct this score, the most severe misbehaviour questions (items 1, 2, 8, 9, 12, 15 and 19) were assigned three points, positive answers of the severe research misbehaviour questions were assigned two points (items 4, 7, 10, 14, 16,18 and 20) and positive answers of the moderate research misbehaviour questions were assigned one point (items 3, 5, 6, 11, 13, 17, 21 and 22). Scores were added up to calculate the composite research misbehaviour severity score (RMSS) (maximum range: 0-43).

Supplementary Table S2. number and percentages per cluster and per determinant. Significance testing by calculating Chi-squares for dichotomous variables and ANOVA for continuous variables such as the H-Index and the PPQ. *We performed a Kruskal Wallis Test for the analysis of the RMSS. RMSS: maximum range 0-43

		Total group (n=535)	Clusters of biomedical scientists			p-value
			Perfectionist (n=140)	Ideal son-in-law (n=193)	Sneaky grandiose (n=202)	
Gender	Male	229	53 (24%)	65 (28%)	111 (48%)	P<0.001
	Female	306	87 (28%)	128 (42%)	91 (30%)	
Age	<40 yr.	396	114 (29%)	141 (36%)	141 (35%)	P<0.05
	>40 yr.	139	26 (19%)	52 (37%)	61 (44%)	
Academic position	PhD student	303	98 (32%)	108 (36%)	97 (32%)	P<0.001
	Postdoctorate, Associate or Assistant Professor	177	38 (21%)	58 (33%)	81(46%)	
	Full Professor	55	4 (7%)	27 (49%)	24 (44%)	
Hirsch index (n=253)		15.4 (IQR 2-24)	10.5 (IQR 2-16.5)	16.6 (IQR 2-26)	16.9 (IQR 3.5-25)	0.058
Research Misbehaviour Severity Score (RMSS)		4.3 (SD+/-4.9)	4.6 (SD +/- 6.1)	3.6 (SD +/- 3.9)	4.9 (SD +/- 5.0)	P<0.05*
Publication Pressure Questionnaire score		42.6 (SD+/-6.6)	43.7 (SD +/- 5.8)	41.4 (SD +/- 6.5)	42.9 (SD +/- 7.2)	P<0.01

Reference List

(1) Ford C.V., Summer G.L. The Disruptive Physician, II: The Role of Personality Disorders. Federation Bulletin 1998; 85:20-29.

(2) G.C.Alexander, J.Humensky, C.Guerrero, H.Park, G.Loewenstein. Brief Report: Physician Narcissism, Ego Threats, and Confidence in the Face of Uncertainty. Journal of Applied Social Psychology 2010; 40:947-955.

(3) Macleod MR, Michie S, Roberts I, Dirnagl U, Chalmers I, Ioannidis JP et al. Biomedical research: increasing value, reducing waste. Lancet 2014; 383:101-104.

(4) Jones DN, Paulhus DL. Introducing the short Dark Triad (SD3): a brief measure of dark personality traits. Assessment 2014; 21:28-41.

(5) Donnellan MB, Oswald FL, Baird BM, Lucas RE. The mini-IPIP scales: tiny-yet-effective measures of the Big Five factors of personality. Psychol Assess 2006; 18:192-203.

(6) Byrne Z.S., Mueller-Hanson R.A., Cardador J.M., Thornton G.C., Schuler H., Frintrup A et al. Measuring achievement motivation: test of equivalency for English, German, and Israeli versions of the achievement motivation inventory. Personality and individual Differences 2004; 37:203-217.

(7) Rosenberg M. Society and the adolescent self-image. Princeton, NJ: Princeton Press; 1965.

(8) Christie R., Geis F. Studies in Machiavellianism. NY: Academic Press; 1970.

(9) Clarkin J.F., Fonagy P., Gabbard G.O. Narcissistic personality disorder. Psychodynamic Psychotherapy for Personality Disorders: A clinical Handbook. 1st ed. Arlington: American Psychiatric Publishing; 2010. p.263.

(10) Skeem JL, Polaschek DL, Patrick CJ, Lilienfeld SO. Psychopathic Personality: Bridging the Gap Between Scientific Evidence and Public Policy. Psychological Science in the Public Interest 2011; 12:95-162.

(11) Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PLoS One 2009; 4:e5738.

(12) Martinson BC, Anderson MS, DeVries R. Scientists behaving badly. Nature 2005; 435:737-738.

(13) Tijdink JK, Verbeke R, Smulders YM. Publication pressure and scientific misconduct in medical scientists. Journal of Empirical Research and Human Research Ethics 2014; 9:64-71.

(14) Tijdink JK, Smulders YM, Vergouwen AC, de Vet HC, Knol DL. The assessment of publication pressure in medical science; validity and reliability of a Publication Pressure Questionnaire (PPQ). Qual Life Res 2014; 23:2055-2062.

(15) Hirsch JE. An index to quantify an individual's scientific research output. Proc Natl Acad Sci 2005; 102:16569-16572.

(16) Cragar DE, Berry DT, Schmitt FA, Fakhoury TA. Cluster analysis of normal personality traits in patients with psychogenic nonepileptic seizures. Epilepsy Behav. 2005;6: 593-600.
(17) Gore PA. Cluster Analysis. In: Tinsley HEA, Brouwn SD, editors. Handbook of Applied Multivariate Statistics and Mathematical Modeling. London: Academic; 2000. 297-321.
(18) Fiske S.T., Gilbert D.T., Lindzey G. Handbook of Social Psychology. Hoboken NJ: Wiley; 2009.
(19) Dolan S.L. Stress, Self-Esteem, Health and Work. Palgrave Macmillan; 2006.
(20) Kornfeld DS. Perspective: research misconduct: the search for a remedy. Acad Med 2012; 87:877-882.
(21) Tijdink JK, Vergouwen AC, Smulders YM. Publication Pressure and Burn Out among Dutch Medical Professors: A Nationwide Survey. PLoS One 2013; 8:e73381.
(22) Fang FC, Casadevall A. Competitive Science: Is Competition Ruining Science? Infect Immun 2015; 83: 1229-33.
(23) Vernon P.A., Villani C.V., VIckers L.C., Harris J.A. A behavioral genetic investigation of the Dark Triad and the Big 5. Personality and individual Differences 2008; 44:445-452.
(24) Cook C, Heath F, Thompson RL. A meta-analysis of response rates in web- or internet-based surveys. Educational and Psychological Measurement 2000; 60:821-836.
(25) Joinson A. Social desirability, anonymity and Internet-based questionnaires. Behav Res Methods Instrum Comput 1999; 31:433-438.

10

Chapter 10

General Discussion

10

Main findings

The aim of the studies described in this thesis was to give a broad description of contemporary biomedical publication culture, focusing on the scientist and trying to identify determinants and consequences of this culture. The conclusions that follow from the data presented in each chapter are:

The Publication Pressure Questionnaire (PPQ) we developed is a valid and reliable instrument to quantify perceived publication pressure among biomedical scientists (chapter 2).

Recently appointed full professors have higher levels of emotional exhaustion and are susceptible to burnout. Furthermore, a relatively high H-index is associated with less burnout symptoms (chapter 3).

Burnout symptoms are associated with perceived publication pressure in biomedical professors. A substantial proportion of them believe that publication pressure has become excessive, and have a cynical view on the validity of biomedical science. These opinions are also correlated to burnout symptoms (chapter 4).

In a sample of 315 Flemish biomedical scientists, no less than 15% admitted they had fabricated, plagiarized or manipulated data in the past 3 years. This self-reported behavior was more common in younger scientists. Furthermore, a composite research misbehavior severity score was strongly associated with perceived publication pressure (Chapter 5).

Focus group interviews suggest that contemporary publication culture leads to negative sentiments, counterproductive stress levels and – arguably most importantly – questionable research practices among both junior and senior biomedical scientists (chapter 6).

This study suggests that psychiatrists fail to account for the effects of industry funding disclosure in their judgement of the reliability and relevance of study results.. On the other hand, psychiatrists are more likely to critically interpret the content of a scientific abstract in which a positive outcome is reported. From our study, it is apparent that there is a striking discrepancy between psychiatrists' attitudes towards the pharmaceutical industry and the effects that funding disclosure has on their perceived credibility and judgement of clinical relevance in study results (chapter 7).

The Machiavellianistic personality trait is associated with self-reported research misbehavior. Higher hierarchical academic positions are associated with higher narcissistic and psychopathic traits, lower self-esteem scores and lower perceived publication pressure. Furthermore, self-reported research misbehavior is more common among those in higher academic positions (Chapter 8).

Biomedical scientists in the 'sneaky grandiose' personality cluster have a relatively high propensity to engage in research misbehavior. A small proportion of the 'sneaky grandiose' might suffer from a psychiatric condition characterized by pathological preoccupation with publishing and being cited. We therefore propose to name this syndrome 'Publiphilia Impactfactorius' (PI), and we suggest this affliction should be considered in revised versions of DSM5 and ICD-10 (Chapter 9)[1].

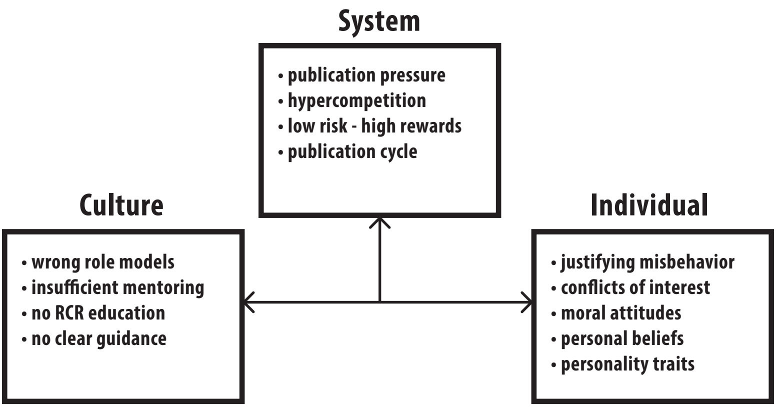
Lessons Learned

Personal reflection

What can we learn from these results? To start with, let's get back to the motivation to start this research project. Initially, from a psychological perspective, I felt that something was out of control in contemporary publication practices. I saw scientists fighting for authorships and battling for being first or last author, and abuse of statistical tests to obtain desired results. The holy grail for most

1 The text was written for the BMJ 2015 Christmas issue and is meant as a 'tongue-in-cheek' article that is based on real data. Unfortunately, it was not accepted for publication after peer-review. It's currently under review by another journal.

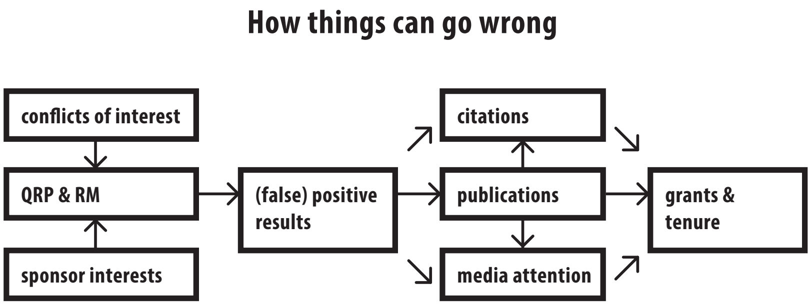
Figure 1. Interacting forces of scientific integrity. Adapted from a PowerPoint slide presented in the keynote lecture of LM Bouter (What is holding us back in the prevention of questionable research practices?) during the 4th World Conference on Research Integrity, Rio de Janeiro, Brazil, 2015



scientists seemed to be publishing in a high impact factor journal. I saw troubling psychological aspects that appeared to sometimes dominate scientific practice: fear of failure, a paranoid attitude towards colleagues, and poor collaboration. In summary, my observations suggested a dominant focus on publishing articles instead of on conducting good and relevant research.

The problem with such personal observations is that they are, by definition, subjective and anecdotal. Nevertheless, I had a sad impression of the present state of biomedical research in the Netherlands, which increased my distrust of research outcomes. Following the dominant academic paradigm, the only way to address this was to investigate whether my impressions were justified and shared by the majority of fellow scientists, or completely biased by my personal beliefs and opinions.

Figure 2. How research misbehavior occur. Adapted from a PowerPoint slide presented in the keynote lecture of LM Bouter (What is holding us back in the prevention of questionable research practices?) during the 4th World Conference on Research Integrity, Rio de Janeiro, Brazil, 2015



Intellectual conflict of interest

Personal beliefs constantly try to entice us to 'bend' our data in a direction that fits our ideas of the world. Researchers naturally become invested in their own hypothesis and in a research process there are continuously opportunities to – unintentionally – disregard data that do not entirely fit the model. Awareness of this confirmation bias was often on top of my head to make sure this thesis would not suffer too much from my prejudices. My opinions, beliefs and perspectives should never trespass scientific scrutiny, methodological rigor and I should make sure that they play at most a minor role in the discussion of the value and generalizability of my work. Due to these biases, modesty about my results is probably appropriate.

Every scientist should be aware of these intellectual and emotional conflicts of interest. It should encourage us to scrutinize our own work, be our most critical reader, be aware of and reflective on our beliefs and personal convictions.

Publication culture

During the research projects reported in this PhD thesis, I could not put off my glasses as a psychiatrist. From a psychiatric perspective, individual incentives for researchers were clearly visible. However, this point of view can narrow your framework of thinking, and my perspective needed to be broadened by other aspects like culture and organizational structure. The interaction of factors from these three domains probably determines where a scientist stands in the spectrum ranging from responsible conduct of research to research misconduct (figure 1). It should be mentioned that even a scientist beholds different positions in this spectrum. His/her morality can take different integrity positions, depending on conflict of interests, prejudices or selfish incentives.(1) These factors encompass the system of science as a whole, the system of publishing and the professional environment with its written and unwritten rules (culture) and individual scientists with their personality traits, emotional states, conflicts of interest, personal believes and moral attitudes.

The scientific culture and the structure within universities (the system) will influence individual scientists profoundly and should not be neglected. The culture includes role models, mentors, education and clear guidance for novel researchers. Important aspects of the system are pressure from institutions to publish, the competitive climate, the reward system and the publication cycle (figure 1). All these factors determine 'how things can go wrong' (figure 2).

These influences can make different individuals behave differently (intentionally and unintentionally) and may lead to responsible conduct of research or to minor or major research misbehavior.

Interpretation of results; what do the results tell us?

A closer look at the results reported in this thesis provides some insight in the aforementioned factors. In our first studies, we investigated the association between perceived publication pressure and phenomena such as emotional exhaustion, burnout and self-reported involvement in questionable research practices. Our research raised new questions regarding the culture and structure of biomedical science. Moreover, the results show a somewhat cynical perspective of Dutch biomedical professors on contemporary publication culture and the credibility of research results. The participating professors felt an excessively high pressure to publish. They express serious doubt on the validity of research results and show high levels of emotional exhaustion, which is the key element of burnout. These results suggest that there are possibly detrimental effects of contemporary publication culture associated with perceived publication pressure. If our academic leaders and role models have cynical thoughts and experience emotional distress, how can they inspire young talented scientists? And to what extent does this influences scientific practices in labs and research institutes?

To gain more insight in these phenomena, we decided to choose a qualitative approach and performed focus group interviews among biomedical researchers in different positions in academia. Our primary goal was to explore what contemporary publication culture means to biomedical scientists, and what problems they report to encounter. The findings picture a predominantly detrimental impact of the present culture on biomedical science and biomedical scientists. The interviewees suggest that often the need for individual recognition and prestige prevails over team effort and collaboration. The participants' responses suggest an overall negative sentiment about the hypercompetitive culture, in which journal impact factors and authorships dominate. They also highlight systemic flaws (ie replication problems due to selective reporting and biased research funding procedures) in contemporary biomedical science. Participants across all hierarchical layers experience high levels of publication pressure and describe vividly the relation of this pressure with questionable research practices. Publication pressure seems to be influenced by several factors (figure 3). This cycle represents adequately the narratives of the participants of the focus group interviews; If you publish, in preferably high impact journals, this will give you more chances of funding allocation to secure future research. With this funding you can hire staff that produce data and if these data 'produce' positive findings, these findings have a high chance for acceptance in high impact journals. This credibility cycle is frequently used by the Dutch organization Science in Transition to pinpoint current publication practices (2).

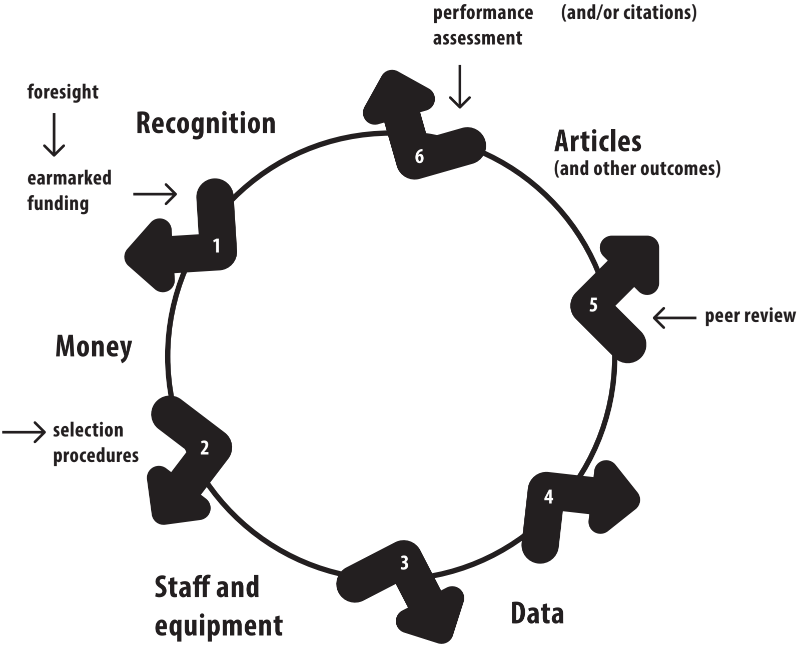
Interestingly, hardly any positive aspects of contemporary publication culture were mentioned. Could it be that many scientists have lost their faith in science or at least in the dominating publishing culture?

A common denominator to our findings is a hypercompetitive climate, wherein research funding, positive results, the number of research papers and authorships, high impact journals, lack of responsible conduct of research, and a chase for status and prestige dominate. Although our study participants view these factors as harmful for the scientific enterprise and expressed a desire to change the present culture, they also feel powerless to do so.

How is it possible that truth finding and objectivity have been partly overtaken by a system of 'publish and perish' that has a number of serious unwanted side effects? It is unlikely that this development was intentional. It seems, however, that the increasing focus on a limited number of quantitative performance indicators may be a threat to core values like trustworthiness, credibility, collaboration, and integrity.

The role of personality traits of scientists has been previously mentioned. Such traits may influence the functioning of scientists and lead to minor or major misbehavior. Narcissistic and egoistic traits have been described (5) as being counterproductive for collaboration and scientific progress, and are anecdotally more common among professors. Chapter 8 concludes that the Machiavellianistic personality trait is associated with self-reported research misbehavior. This finding should be interpreted with caution; survey statistics and the cross sectional design have major limitations that make identification of a causal relation troublesome. However, it seems logical that people with a Machiavellianistic character ('a tendency to be unemotional, detached from conventional morality and hence to deceive and manipulate others, focused on unmitigated achievement and high priority of own performances' (6)) have a higher propensity towards (research) misbehavior. This trait was also associated with higher academic positions hypothesizing that this trait will probably help you to get ahead in your scientific career.

Figure 3. The credibility cycle, adapted from Latour and Woolgar (1986). Points at which organizational devices connect to the cycle are shown. (3;4)



Scientific integrity

A healthy scientific culture is one in which research is conducted in a responsible and methodologically sound way, with mentally healthy researchers in an ambitious and mildly competitive environment. The opposite of responsible conduct of research is research misbehavior, often divided into: 1) 'scientific misconduct' (such as fabrication, falsification and plagiarism, and 2) less severe breaches of research integrity; the 'questionable research practices'(7).

Our studies suggest that not only the research system (in terms of perceived publication pressure) influences scientific behavior (chapter 5), but that also the personality of the researcher can 'provoke' unwanted research behavior (chapter 8). These findings shed new light on putative causes of research misbehavior. Before, fraudsters were often seen as 'just a few bad apples' among generally responsibly behaving researchers. Recent evidence, however, suggests that these misbehaviors are far more common (8). Factors such as perceived publication pressure and personality traits should be taken into account when analyzing the causes of research misbehavior. Furthermore, evidence suggests that dishonesty and consequently possible breaches of research integrity due to conflict of interest situations should not be solely regarded as good versus bad. Dishonest actions are also 'committed' by persons with high moral standards, who think highly of themselves in terms of honesty and role modelling, when they face a complex situation. When they are tempted to follow selfish incentives at the expense of stretching integrity rules, they will typically consider these breaches morally acceptable and find 'valid' reasons for these actions. (1)

Governmental policy makers, leaders of universities and research institutes, and heads of departments should probably allow for personality factors, not only in their application procedures of hiring talented scientists, but also in their educational and prevention programs on research integrity.

Is competition ruining science?

To identify possible causes of this somewhat negative and slightly hostile environment, let's get back to the role of competition in science as described in the introduction of this thesis. Although we know that competition has positive and stimulating effects on people and cultures (9;10), negative consequences can outweigh possible beneficial effects, certainly when competition evolves into 'hypercompetition'. Our studies partly confirm earlier research that states that hypercompetition is backfiring on the academic enterprise. (11) The results in chapter 6 clearly point towards hypercompetition. Many participants of the focus groups describe hypercompetition causing counterproductive pressure. The 'rat race' for grants, authorships, etc. clearly affects them, and is dominated by high achievers that will be rewarded with high impact papers, new grants and in the long run a tenured position. Participants show some despair in this respect. Publication pressure is seen as a consequence of hypercompetition; most biomedical scientists will acknowledge that the number of publications, and the proportion of high impact factor publications among them, are of overwhelming importance for a successful scientific career.

Compared to other chapters, chapter 7 is a stranger in our midst. The link with the other studies is however quite interesting. The other chapters focus on the role of publishing scientific material in the broadest sense. In Chapter 7 we look at how psychiatrists judge a (industry sponsored) scientific abstract. Their judgement was not influenced by industry funding disclosure in the abstract. Although we did not study whether this effect influences medical decision making, we should allow for the enormous influence funding can have on the outcome of study results. Although psychiatrists discounted for this bias by perceiving a negative study as more credible, psychiatrists and physicians in general should probably be more aware of this potential bias.

Consequences

When publication pressure is at a healthy level, scholars get the sunny side of a liberal publish or perish culture that rewards their efforts. But if the pressure is too strong, scientists feel stressed, function suboptimal and report an association with minor and major scientific misconduct. (12)

Another frequently expressed sentiment is that biomedical scientists feel that the content of publications seems to be less and less important in academic medicine, and that the new rule is that 'it no longer matters what you write, but only how often, where and with whom you publish' (13). Modern science has adopted a culture of ranking and evaluation in which citations, publications and other measurable output strongly dominate. This sentiment was also expressed in our focus group interviews across all academic ranks, as summarized in Chapter 6.

The issue is how the modern research culture influences individual behavior. If researchers are under too much pressure would they be more willing to involve in questionable research practices (or worse) to get ahead? Our research unfortunately seems to confirm this notion. To further speculate, this pressure most likely negatively influences other academic important duties, such as teaching, mentoring, and collaboration.

Methodological considerations

Methodological strengths

We used different web-based survey methods to study biomedical researchers anonymously on a specially protected website. Good survey methods include the use of validated questions and questionnaires, clearly formulated research questions, accurate data and advanced nonresponse follow-up techniques. These were all used and applied in our surveys (chapters 3, 4, 5 and 8).

To give more insight in correlations between the different validated questionnaires used in chapter 3,4, 8 and 9, we have added two tables as appendixes to this general discussion (see the end of this chapter).

We used exploratory and confirmatory item response theory to investigate whether our publication pressure questionnaire was a valid and reliable instrument to measure publication pressure (chapter 2).

Since data on publication culture is scarce, we decided to do exploratory, hypothesis-generating qualitative research by conducting focus group interviews. Qualitative research methods are valuable to determine perceptions of a culture, generated by open research questions (14;15). An ideal combination is to collect insight in-depth information by a qualitative approach and confirm the results and findings in a quantitative approach.

Major limitations

A number of limitations must be acknowledged. First of all, survey techniques have intrinsic shortcomings. Most surveys have a substantial percentage of non-responders, which will most likely cause response bias. The question is in which direction response bias will distort the findings. This is often hard to determine, and it is often conceivable that response bias occurs in opposite directions. A good example is given in chapter 4. We think that response bias in this study may have been bidirectional; on the one hand, non-response may be related to lack of time or sense of task overload. Non-response in our survey may have caused underestimation of burnout symptoms and of suffering from perceived publication pressure. On the other hand, lack of cooperation may also be related to the subject of the survey. Possibly, some medical professors consider publication pressure to be irrelevant, and refuse to participate. Such bias would conceivably have caused overestimation of burnout and disapproval of publication pressure among respondents.

Secondly, for advanced survey techniques you need to know what factors determine a good response rate. For this knowledge, experience is essential. Four studies described in this thesis used surveys to collect the data in this thesis has helped us to improve our survey techniques due to more specific information of the participants, better insight in active email addresses and information whether the email was read by the participant. This resulted in higher response rates (in chapter 7 and 8 response rates of +/- 50%). This has helped to diminish response bias.

Thirdly, the operationalization of perceived publication pressure is difficult and subject to controversy. The validated questionnaire of the PPQ consisted of 14 items. The question is whether these items really represent all relevant aspects of a perceived publication pressure. To prepare the perfect conditions for an ideal construct validation ideally a gold standard is needed that should be used as a reference for the new scale that is being validated. Since such scale did not exist, we used the MBI to have a validated measure that can be used as a reference point. Besides, it could be questioned whether the MBI can be related to the PPQ and at the same time being used as a reference point in the validation study of Chapter 2.

The focus group interviews pointed to other important aspects of publication pressure (such as the need of getting funding and having positive findings) that were not included of the 14-item PPQ.

Fourth, surveys with a cross sectional design cannot prove causal relations and should therefore always be interpreted with caution. The better option would be a longitudinal cohort design in which the putative determinants would be measured before the self-reported misbehaviors.

Fifth, self-reported misbehaviors have limited validity and reliability. They are subjective and vulnerable to cheating, including self-deception. Scientists may not easily report their own misbehavior and have most likely not mentioned all their misbehaviors, even when anonymity is guaranteed. This might have resulted in underestimation of the actual prevalence of these misbehaviors. Furthermore, we have related the personality traits with an ad hoc composite score of the research misbehaviors. This score we did not validate and the weighing of the 22 items is somewhat arbitrary.

Towards a more sound academic enterprise

The findings contained in this PhD thesis have provided some new insights. First, they add evidence to the upcoming field of research integrity. Recent fraud cases have created increased awareness and attention towards research culture and research misbehavior. Future research must create a balanced view on the integrity problems science is facing today. That research will probably show that, on the aggregate, the aggregated impact of the common minor research misbehaviors is much larger than that of the rare instances of fabrication and falsification. This can eventually lead to new paths for prevention of research misbehavior and for fostering responsible conduct of research.

Although firm conclusions on the causes of research misbehavior are difficult to draw from the results in this thesis, we would like to go beyond the scope of the thesis and deliberately try to picture the evolution the scientific enterprise can – and should – consider.

Policy makers and scientific leaders should aim at changes on a systemic and individual level. Structural changes should address hypercompetition and create a more balanced reward system (16) for scientific achievements; set common goals, stimulate collaboration, reduce the importance of the impact factor (17) and unselfishness should be recognized, valued and rewarded.

An individual approach should encompass personality characteristics and traits in biomedical scientists, and develop prevention strategies. An institution can discourage Machiavellianistic and selfish behavior by stimulating collaboration and reducing the focus on individual performance. Talented scientists should receive stable support to enable them to focus fully on their research and not be too much bothered by financial stress. Creativity thrives on freedom and interactivity and thus asks for rewards for collaboration. It's time for the scientific enterprise to evolve and set new goals for scientific values in a more stimulating atmosphere.

Second, our findings give support to recent pleas for 'slow science'. (18;19) More specifically our results suggest the need to discourage individualism, Machiavellianism, narcissism and 'publiphilia', and to let scientific rigor and quality prevail above impact factors and other citation scores. Although alternative quantitative measures to rank scientists are sparse and lacking firm evidence, and 'soft' measurements such as scientific quality are extremely hard to rank, policy makers should consider changing the ranking system. This can be achieved by ranking scientists not on their citation records but rank scientists on other measurements such as the softer methodological sound aspects of their work and alternative metrics.

Third, the scale of the academic scientific enterprise is enormous and still increasing; more young scientists join academia every year and they produce more and more papers, producing the next generation of scientists during the process. However, there is not much space in terms of tenured positions, and available research grants. We may have to decide to reduce the scale: less scientists and less papers. This implies there should be a more restrictive appointment policy of scientists, especially in the biomedical field. In academic medicine in the Netherlands, a PhD is often obtained by a medical doctor who is not trained for conducting research nor particularly motivated to conduct research. The question is whether it is a sound idea for so many young physicians who aspire to become medical specialists to first write a PhD thesis. It is somewhat strange that a PhD degree became an important entry criterion for medical specialization. Science can, in principle, improve medical doctors, but vice versa, medical doctors don't necessarily make science better.

Fourth, fostering creativity and originality for new research ideas is crucial. The role of creativity and originality in the scientific process is – in the exploration phase – often suggested as a key element in conducting innovative research. One has to think creative to find new solutions for scientific problems. The effect of a hypercompetitive culture on research creativity is unknown. Creativity is seen as a characteristic that results from intrinsic rather than extrinsic motivation, and requires sufficient incubation time. (20) Creativity opens up when an individual passionately chases ideas. A creative individual feels like playing rather than working (21). Creativity is encouraged when scientists face challenging problems; have autonomy to seek own solutions, and act in collaboration-friendly settings where they can easily network with others (22) Please note that while creativity is essential for scientists when in the explorative mode, in the justification phase of research projects, creativity is not the essential ingredient and even a potential risk; in this phase, a scientists should be concise, honest and stick to the study protocol and the analysis plan.

People who are influenced and motivated by external rewards, such as in present biomedical research culture (high impact factor publications, grants, tenure), are less likely to explore 'creative solutions'. In the process of being creative, a reasonable amount of job security and sense of independence seems to be important in fostering creativity. (21) Although creativity may in fact require a certain degree of competition, hypercompetition may backfire on creativity. These effects include the fear of employment loss faced by many scientists today (23).

Fifth, there is an overload of scientific assessment by metrics. Citations, journal impact factors, Shanghai Index for universities, h-index, Twitter and even Facebook likes. There are numerous quantitative measures that are shaped to measure science and scientific output. But how robust and reliable are these metrics to assess scientific quality and what importance should we give them in the future assessment of science and scientists. During the last decade, metrics have become dominant in the assessment of scientists, institutions and scientific quality. The use of a small number of quantitative assessment criteria can certainly become a perverse incentive and distort the scientific process. Poorly designed evaluation criteria are "dominating minds, distorting behavior and determining careers." (24) According to a recent review performed by the UK-research council (25). This review makes 20 specific recommendations (ie nuanced use of metrics by institutions and individual scientists) for further work and action by stakeholders across the UK research system. These recommendations should be seen as part of broader attempts to strengthen research governance, management and assessment (25). From Dutch soil, Science in Transition has questioned the overreliance on a few metrics in the Dutch research system and have managed to reform assessment structures in evaluation protocols (26;27)

Taken together, the 'tournament structure' (9) of science probably ought to become less dominant. This will create a more sound and more healthy moderately competitive environment with fewer scientists.

Do biomedical scientists desire change?

An interesting observation among participants of the focus group interviews was the outspoken desire to change the academic system with its systemic flaws. Nevertheless, their feeling that they are unable to change the present academic enterprise, withheld them to actually start reforming contemporary culture and structure. This suggests that they are not feeling in control.

Future research

The scarce literature available is predominantly based on personal beliefs and opinions. The available empirical evidence consists mostly of qualitative research or survey data that describe cultural aspects of academia. This thesis is part of that tradition. Most biomedical scientists consider this type of 'evidence' of inferior value. That is probably too harsh, but there is definitely a need for experiments on the efficacy of systemic, cultural and educational interventions. Although thorough understanding of publication practices is clearly lacking, the qualitative approach can point out the fundamental pitfalls and challenges that contemporary science encounters. An in-depth interview study with heads of departments, institutional leaders and governmental policymakers should uncover more insight in their thoughts about the current publication culture, their hesitations to intervene, and may address conceivable solutions.

For future research on research misbehavior, a focus on identifying its determinants may shed more light on possible interventions and can lead to detection of persons at risk for misbehaving during the early stage of research careers. This can help to develop a model to predict research misbehavior on a personal and on a structural level. More insight in the personality of scientists cutting corners or worse is necessary to know how interventions on a structural level will result in prevention. This could for instance be done in a prospective cohort study based on the participants of chapter 8. We could conduct a second survey in 2017 and compare these findings and traits with the earlier results to determine whether personality traits and/or academic rank are a risk factor for self-reported research misbehavior.

In the data used in chapter 8, we can also analyze 'determinants' for responsible conduct of research. Scientists that have low scores on our composite score of self-reported research misbehavior might have a personality profile that can be protective against unwanted behaviors. Uncovering these determinants can help to find the best reward strategies for good scientific behavior.

Another issue is the selection procedure in academia. Should personality trait assessment be part of routine procedure in selection for the top positions in academia? Should this also be applied to funding allocation? Is it possible that academic medicine automatically selects the more narcissistic and Machiavellianistic personalities for their top positions, thus becoming more prone to research misbehavior? Future in-depth qualitative research on selection procedures can surely shed more light on this important subject. This can be done by observational studies that describe the psychology of tenure allocation and professorships. Finally, an intervention study should focus on how researchers work by letting them be more reflective when they are facing ethical and integrity difficulties throughout the research process. A kind reminder on the ethical research code of an institute before a scientist submit a paper might influence moral decision making in conducting research.(1) A randomized two armed study design can determine whether researchers are more willing to honor the research code and would be enticed for responsible conduct of research.

In conclusion, this thesis draws a somewhat dark picture of the academic enterprise. We describe a culture of high perceived pressure to publish, with cynical views on the validity of research results. Possibly due to this pressure some biomedical scientists are emotionally exhausted and some report to engage in minor or even major research misbehavior. This self-reported misbehavior is associated with specific personality traits that are less compatible with ethical values such as honesty, conscientiousness, integrity and trustworthiness.

Are these our findings only sketching a sorrowful picture? No, the results also give hope and perspective. Many of the problems are challenging. Change starts with reflection and awareness. This awareness is rapidly increasing within the scientific community, which already leads to intervention strategies – albeit on a small and experimental scale - that will help not only policy makers and heads of departments, but first and foremost colleagues and co-scientists.

Appendix A

Table 1. Correlation matrix of scales used in Chapter 3 and 4

* p<0.01

** p<0.05

Total PPQ		Total PPQ	UBOS EE	DP	PA	UBES Vigour	Dedication	Absorption
Utrecht BurnOut Scale (UBOS)	Emotional Exhaustion (EE)	1						
	Depersonalisation (DP)	.45**	1					
	Personal Accomplishment (PA)	.29**	.58**	1				
Utrecht BEvlogenheids Schaal (UBES)	Vigour	-.15**	-.28**	-.31**	1			
	Dedication	-.18**	-.36**	-.27**	.61**	1		
	Absorption	-.24**	-.38**	-.35**	.56**	.80**	1	
		-.04	-.07	-.12**	.45**	.72**	.68**	1

Table 2. Correlation matrix of scales used in Chapter 8 and 9

* p<0.01

** p<0.05

	Big5Extra versionTotal	Big5Agreeable nessTotal	Big5Concious nessTotal	Big5neuro ticismTotal	Big5Intel lectTotal	Total PPQ	Achievement Motivation Scale	SelfEsteem ScaleRosen berg	Narcism9 itemTotal Score	Machiavel lism9item TotalScore	Psychopathy 9itemTotal Score	Cynism TotalScore	QRP_bewerkt SumScore
Big5Extra versionTotal	1												
Big5Agreeable nessTotal	.313**	1											
Big5Concious nessTotal	-.104*	.086*	1										
Big5neuro ticismTotal	-.115**	.023	.008	1									
Big5Intel lectTotal	.203**	.169**	-.137**	.080	1								
TotalPPQ	-.044	-.102*	.015	.244**	.012	1							
Achievement MotivationScale	-.168**	-.077	-.309**	.086*	-.032	.016	1						
SelfEsteemScale Rosenberg	-.231**	.073	-.140**	.481**	-.023	.055	.164**	1					
Narcism9item TotalScore	.367**	-.056	-.007	-.081	.125**	.013	-.282**	-.343**	1				
Machiavellism 9itemTotalScore	-.097*	-.235**	.025	.118**	.021	.091*	-.018	.040	.241**	1			
Psychopathy 9itemTotalScore	.118**	-.184**	-.146**	.097*	.010	-.017	.042	.048	.373**	.417**	1		
CynismTotalScore	.160**	.212**	-.027	-.180**	.045	-.176**	-.032	-.191**	-.135**	-.491**	-.324**	1	
QRP_bewerkt SumScore	.024	-.124**	-.069	-.013	.009	.112**	.001	.000	.086*	.135**	.159**	-.062	1

10

Reference List

(1) Mazar NF, Ariely D. Dishonesty in scientific research. J Clin Invest 2015;125:3993-6.

(2) Science in Transition. www.scienceintransition.nl. 2015.

(3) Hessels L.K., van Lente H., Smits R. In search of relevance: The changing contract between science and society. Science and Public Policy 2009;36:387-401.

(4) Latour B., Woolgar S. Laboratory Life. Princeton: Princeton University Press; 1986.

(5) O'Boyle EH, Forsyth DR, Banks GC, McDaniel MA. A meta-analysis of the Dark Triad and work behavior: a social exchange perspective. J Appl Psychol 2012;97:557-79.

(6) Christie R., Geis F. Studies in Machiavellianism. NY: Academic Press; 1970.

(7) Steneck NH. Fostering integrity in research: definitions, current knowledge, and future directions. Science and Engineering Ethics 2006;12:53-74.

(8) Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. PLoS One 2009;4:e5738.

(9) Fang FC, Casadevall A. Competitive Science: Is Competition Ruining Science? Infect Immun 2015;83:1229-33.

(10) Merton R.K. Priorities in scientific discovery. Am Sociol Rev 1957;22:635-59.

(11) Alberts B, Kirschner MW, Tilghman S, Varmus H. Rescuing US biomedical research from its systemic flaws. Proc Natl Acad Sci U S A 2014;111:5773-7.

(12) van Dalen HP, Henkens K. Intended and unintended consequences of a publish-or-perish culture: A worldwide survey. Journal of the American Society for Information Science and Technology 2012;63:1282-93.

(13) Fang FC, Casadevall A. Competitive Science: Is Competition Ruining Science? Infect Immun 2015;83:1229-33.

(14) Creswell J.W. Qualitative inquiry and research design: choosing among five traditions. Thousand Oaks: Sage; 1998.

(15) Kitzinger J. Qualitative research. Introducing focus groups. BMJ 1995;311:299-302.

(16) Ioannidis JP, Khoury MJ. Assessing value in biomedical research: the PQRST of appraisal and reward. JAMA 2014;312:483-4.

(17) Casadevall A, Fang FC. Causes for the persistence of impact factor mania. MBio 2014;5:e00064-14.

(18) The Slow Science Manifesto. http://slow-science.org/. 2015.

(19) Owens B. Long-term research: Slow science. Nature 2013;495:300-3.

(20) Hennessey BA, Amabile TM. Creativity. Annu Rev Psychol 2010;61:569-98.

(21) Amabile TM. Creativity in context. Boulder, CO: Westwood Press; 1996.

(22) Alon U. How to build a motivated research group. Mol Cell 2010;37:151-2.

(23) Rick S, Loewenstein G. Hypermotivation (Commentary on 'Mazar et al., The dishonesty of honest people.'). J Market Res 2008;45:645-8.

(24) Lawrence PA. The mismeasurement of science. Curr Biol 2007 Aug 7;17(15):R583-R585.
(25) Wilsdon J.et al. The Metric Tide: Report of the Independent Review of the Role of Metrics in Research Assessmentand Management. 2015.
(26) Standard Evaluation Protocol 2015. https://www.knaw.nl/nl/actueel/publicaties/standard-evaluation-protocol-2015-2021 . 2016.
Ref Type: Online Source
(27) Science in Transition, position paper. http://www.scienceintransition.nl/over-science-in-transition/position-paper . 2016.
Ref Type: Online Source

11

Chapter 11

English Summary

The aim of the studies described in this thesis was to give a broad description of contemporary biomedical publication culture in the Netherlands, with a special focus on the individual scientist. We have tried to identify both the determinants and the consequences of this culture. The conclusions that follow from the interpretation of the data can be divided in three main subjects; publication pressure, research misbehavior and publication culture.

Publication pressure

Publication pressure is the pressure that scientists perceive to publish the results of their research in scientific journals. We developed a questionnaire to measure perceived publication pressure, the Publication Pressure Questionnaire (PPQ). We analyzed the psychometric properties of the questionnaire with confirmatory and explanatory factor analyses. The fit (of each subscale) was then further investigated by means of item response theory (IRT). We concluded that the PPQ is a valid and reliable instrument to quantify perceived publication pressure among biomedical scientists (Chapter 2).

We have sent a web-based survey to 437 Dutch full professors to study publication pressure with the PPQ and burn out with the Utrecht BurnOut Scale (UBOS). We concluded that publication pressure is strongly correlated with emotional exhaustion (Chapter 4). Also, recently appointed full professors have higher levels of emotional exhaustion and are thus particularly susceptible to burnout. Furthermore, a relatively high H-index was associated with less burnout symptoms (Chapter 3). A substantial proportion of professors believed that publication pressure has become excessive, and expressed a cynical view on the validity of biomedical science.

Research misbehavior

Research misbehavior can be defined as different dishonest behaviors in the scientific process that may corrupt findings and conclusions. Fabrication, falsification, and plagiarism are usually qualified as research misbehavior , although it can be questioned whether the latter should be considered as such. Other actions may be referred to as questionable research practices (QRP's), typical examples of which include salami slicing, gift authorships or 'intuitively' deleting data.

Particularly in medicine, concerns have been expressed about the high prevalence of research misbehaviors. In 2009, a systematic review concluded that almost 2% of scientists confessed having fabricated or falsified data at least once during the last 3 years, and up to 33% admitted to other questionable research practices (Fanelli 2009).

We studied biomedical scientists to determine correlates of self-reported research misbehaviors. In a cross-sectional study we sent out a survey with different questionnaires to two different study populations. In a sample of 315 Flemish biomedical scientists, no less than 15% admitted they had fabricated, plagiarized or manipulated data in the last 3 years. This self-reported behavior was more common in younger scientists. Furthermore, a composite research misbehavior severity score was strongly associated with perceived publication pressure (Chapter 5).

In another cross-sectional study of 535 Dutch biomedical researchers we correlated personality traits with self-reported research misbehavior. The Machiavellianistic personality trait was associated with research misbehavior. Furthermore, higher hierarchical academic positions were associated with higher narcissistic and psychopathic traits, lower self-esteem scores and lower perceived publication pressure. Self-reported research misbehavior was more common among those in higher academic positions (Chapter 8).

Publication culture

It is very difficult to determine what publication culture exactly is. We have explored the definition by asking almost 100 Dutch biomedical scientists among all academic ranks to hypothesize which themes should be part of the definition of publication culture. In a focus group study, participating scientists suggested key aspects of the contemporary publication culture are publication bias, authorships, funding, hypercompetition and citation indexes such as the Impact Factor. The themes discussed in these interviews were mainly perceived as negative. They can lead to cynical sentiments, counterproductive stress levels and - arguably most importantly - questionable research practices among both junior and senior biomedical scientists (Chapter 6).

Two of the themes that were mentioned in the focus group interviews were positive outcome bias and the influence of the pharmaceutical companies on research results. This bias was studied in chapter 7 by cross-sectional survey among 400 Dutch psychiatrists. The results suggested that psychiatrists fail to account for the effects of industry funding disclosure in their judgement of the reliability and relevance of study results. On the other hand, psychiatrists are more likely to critically interpret the content of a scientific abstract in which a positive outcome is reported. That is a striking discrepancy between psychiatrists' attitudes towards the pharmaceutical industry and the effects that funding disclosure has on their perceived credibility and judgement of clinical relevance in study results.

Appendix

List of Publications

Publications resulting from this thesis

Tijdink JK, Vergouwen AC, Smulders YM. Publication pressure and burn out among Dutch medical professors: a nationwide survey. PLoS One 2013;8:e73381.

Tijdink JK, Smulders YM, Vergouwen AC, de Vet HC, Knol DL. The assessment of publication pressure in medical science; validity and reliability of a Publication Pressure Questionnaire (PPQ). Qual Life Res 2014;23:2055-62.

Tijdink JK, Vergouwen AC, Smulders YM. Emotional exhaustion and burnout among medical professors; a nationwide survey. BMC Med Educ 2014;14:183.

Tijdink JK, Verbeke R, Smulders YM. Publication pressure and scientific misconduct in medical scientists. J Empir Res Hum Res Ethics 2014;9:64-71.
Tijdink JK, Schipper K, Bouter LM, Maclaine Pont P, De Jonge J, Smulders YM. How do scientists perceive the current publication culture? A qualitative focus group interview study among Dutch biomedical researchers. BMJ Open 2016;6:e008681

Tijdink JK, Smulders YM, Bouter LM, Vinkers CH. How does industry funding disclosure influence psychiatrists? A randomized trial among Dutch psychiatrists. 2015. Under review.

Tijdink JK, Bouter LM, Veldkamp C, van de Ven P, Wicherts J, Smulders YM. Machiavellianism is associated with research misbehavior in Dutch biomedical scientists. 2015. Under review.

Tijdink JK, Smulders YM, Bouter LM. Publiphilia Impactfactorius: a new psychiatric syndrome among biomedical scientists? 2015. Submitted.

+

Other Publications

Publications related to the subject of this thesis

Tijdink JK, Vergouwen AC, Smulders YM. The happy scientist. Ned Tijdschr Geneeskd. 2012;156:A5715. Dutch.

Tijdink JK, de Rijcke S, Vinkers CH, Smulders YM, Wouters P. Publication pressure and citation stress; the influence of achievement indicators on scientific practice. Ned Tijdschr Geneeskd. 2014;158:A7147. Review. Dutch.

YM Smulders, JK Tijdink. Unreliable science (or scientists)? AMSj Februari 2015.

Vinkers CH, Tijdink JK, Otte WM. The use of positive and negative words in scientific abstracts over time: a retrospective analysis of PubMed abstracts between 1974 and 2014. BMJ. 2015;351:h6467.

Peer reviewed publications not related to the subject of this thesis

Tijdink JK, Soethout MB, Koerselman GF, ten Cate TJ. The interest shown by medical students and recently qualified doctors in a career in psychiatry. Tijdschr Psychiatr. 2008;50:9-17. Dutch.

Sijl AM van, Tijdink JK, Vergouwen ACM. Antipsychotica bij katatonie. Tijdschrift voor Neuropsychiatrie & Gedragsneurologie 2009;8:69-72

Tijdink JK, van den Heuvel J, Vasbinder EC, van de Ven PM, Honig A. Does on-site urine toxicology screening have an added diagnostic value in psychiatric referrals in an emergency setting? Gen Hosp Psychiatry. 2011 Nov-Dec;33(6):626-30. doi: 10.1016/j.genhosppsych.2011.07.008

Vinkers CH, Tijdink JK, Luykx JJ, Vis R. Choosing the correct benzodiazepine: mechanism of action and pharmacokinetics. Ned Tijdschr Geneeskd. 2012;155:A4900. Review. Dutch.

Luykx JJ, Vis R, Tijdink JK, Dirckx M, Van Hecke J, Vinkers CH. Psychotic symptoms after combined metronidazole-disulfiram use. J Clin Psychopharmacol. 2013;33:136-7. doi: 10.1097/01.jcp.0000426185.68487.9a.

De Kort S, Tijdink JK. Verleiden tot terughoudendheid. Podium voor Bio-ethiek. 2014;21:18-21

Vinkers CH, van de Kraats GB, Biesaart M, Tijdink JK. Is my patient able to provide informed consent? A practical guideline. Ned Tijdschr Geneeskd. 2014;158:A7229. Review. Dutch.

Djelantik M, Bloemkolk D, Tijdink JK. Wernicke encephalopathy. Tijdschr Psychiatr. 2015;57:210-4. Dutch.

Tijdink JK, Katzenbauer M. Onvoltooid en ongepubliceerd onderzoek: verspilling en vertekening. Tijdschrift voor Psychiatrie 57 (2015);3:225 - 225

Tijdink JK, Smulders YM, Biesaart MC, Vinkers CH. Suicidality and alcohol abuse. Ned Tijdschr Geneeskd. 2015;159:A9159. Dutch.

Tijdink JK. Psychiaters zijn zo gek nog niet. Essay over psychiaters en psychiatrie. 2015. Uitgeverij Diagnosis.

Katzenbauer M. Tijdink JK. Genetische profielen die zijn geassocieerd met schizofrenie en bipolaire stoornis voorspellen creativiteit nauwelijks Tijdschrift voor Psychiatrie 57 (2015);11:846 - 846

Dankwoord

Dankwoord

Vóórdat ik startte met het schrijven van dit dankwoord wist ik helemaal niet hoe ik een dankwoord moest schrijven en wat ik in het dankwoord moest zetten. Natuurlijk, ik ben dankbaar. Dankbaar voor de hulp, adviezen, deelname aan mijn onderzoeken, 'spiritual guidance' en reflecties. Dankbaar voor het plezier wat dit proefschrift me heeft gebracht. En niet in de laatste plaats dankbaar voor de ontmoetingen met inspirerende personen die me aan het denken hebben gezet. Zij waren van onschatbare waarde bij de totstandkoming van dit hoofdstuk. Maar hoe verwoord je dat in een dankwoord?

Om deze traditie na te leven en om niemand voor het hoofd te stoten heb ik een aantal proefschriften opengeslagen om te ontcijferen wie je precies moet bedanken en in welke volgorde. De meeste dankwoorden beginnen bij het bedanken van de patiënten, dan de promotoren, eventuele copromotoren, medeauteurs, collega onderzoekers, kamergenoten, andere collegae die je hebben geïnspireerd. Nog meer collega's die je tijdens de laatste loodjes 'erdoorheen' hebben geholpen. Daarna worden je vrienden bedankt, de paranimfen, dan je ouders (en in bijzondere gevallen je schoonouders) en tot slot de liefde van je leven (met eventueel de bijbehorende kinderen die zijn ontstaan tijdens het langdurige promotietraject). Om u een beetje te verassen; die volgorde heb ik omgedraaid.

Wat me tevens opviel aan het bestuderen van de verschillende dankwoorden was dat er veel ongeschreven hoffelijkheden zijn waaraan een dankwoord voldoet. Ik heb geprobeerd om in mijn dankwoord aan die ongeschreven regels te voldoen en heb geprobeerd deze regels samen te vatten in 7 adviezen, hopelijk behulpzaam voor toekomstige promovendi:

1. Wees niet te eerlijk in je dankwoord. Woorden kunnen verkeerd geïnterpreteerd worden. Daarnaast, vanwege een selectief geheugen vergeten mensen nu eenmaal de positieve woorden, en blijven negatieve woorden vaak en langer doordreunen. Dat is zonde.
2. Vergeet niet je vijanden te bedanken. Iedere wetenschapper heeft nu eenmaal vijanden. Ze zorgen ervoor dat je scherp blijft en je je opvreet over nutteloze zaken waar je je niet over wil opvreten. Toch zorgen ze ervoor dat je keer op keer gemotiveerd raakt om te groeien en beter te presteren.
3. Wees lief voor je collega's. Je komt ze nu eenmaal in de toekomst weer tegen en lieve woorden helpen je verder in het leven, als mens en als wetenschapper.
4. Dankwoorden bevatten clichés en dat komt omdat clichés meestal de waarheid bevatten. Als u anti-clichés bent – en dat is op zich ook al een cliché - raad ik u aan om u over uw antipathie te zetten en wel clichés in uw dankwoord op te nemen.
5. Bedank wie er gevoelig is voor afwijzing of krenking. Juist die collega's of vrienden die wat gevoeliger zijn willen graag genoemd worden in een dankwoord.
6. Bedenk goed naar wie je het proefschrift stuurt en controleer of iedereen die het ontvangt in het dankwoord staat. Simpele en doeltreffende methode.
7. Benoem de mensen die je vergeten bent te bedanken. Een standaard alinea kan zijn: 'Dank iedereen die ik vergeten ben te benoemen in het dankwoord. Er zit geen intentie achter, het is onbewust en dat ik u vergeten ben is mijn fout. Het wordt mogelijk veroorzaakt door een selectief geheugen, door selectiebias, door promotiestress of vanwege het feit dat u onvoldoende indruk heeft gemaakt om u in het dankwoord op te nemen.'

+

Het officiële dankwoord

Liefste Roos, mijn dankbaarheid naar jou heeft maar weinig met mijn proefschrift te maken. Het heeft eerder met het leven te maken. Desalniettemin zijn de momenten schaars om je publiekelijk te bedanken. En dat is jammer want eigenlijk verdien jij pagina's lange epistels. Een proefschrift is dan misschien maar een schamele plek om een keer mijn dankbaarheid voor dat je bent te verwoorden. Daar gaat ie: voor je tomeloze, hartverwarmende, oeverloze, overtuigende en onbevangen liefde ben ik je dankbaar. Voor je schoonheid, je vrouwelijkheid, je elegantie en je gracieus zijn ben ik je dankbaar. Voor je inzicht en je geduld ben ik je dankbaar. Voor je wijze woorden die mij rustig en onrustig maken ben ik je dankbaar. Voor alle grote en kleine dingen ben ik je dankbaar. Voor de manier hoe jij moeder bent van onze kinderen ben ik je dankbaar (ik zou iedereen zo'n geweldige moeder wensen). Voor de manier hoe jij soms hysterisch kan zijn ben ik je dankbaar. Voor de manier hoe je mij accepteert ben ik je dankbaar. Voor de manier hoe jij me daadkrachtig maakt en me stimuleert om alles wat nieuw is te onderzoeken ben ik je dankbaar. Voor de manier hoe je perfect bent ben ik je dankbaar. Ik hoop dat ik nog lang dicht bij je mag blijven.

Lieve Evy (en toekomstige kinderen), volgens de ongeschreven richtlijn hoor jij ook in het dankwoord. Alle clichés over kinderen zijn waar. Ze zijn een spiegel voor mijn eigen ziel en wijzen me keer op keer weer op mijn eigen onvolkomenheden, kwetsbaarheden en inconsistenties. Evy, je maakt me wholehearted.

Lieve pap en mam, zonder jullie was dit proefschrift er niet eens geweest. Ook al hebben jullie niets direct bijgedragen, zonder jullie onvoorwaardelijke liefde en gehechtheid was er niets van me terecht gekomen. Daarvoor is alleen maar nederigheid en dankbaarheid gepast. Nederigheid en dankbaarheid voor het immense kracht die daarvan uitgaat.

Lieve vrienden, jullie zorgen er keer op keer weer voor dat ik geïnspireerd raak, nieuwe ideeën krijg en nooit zal opgeven. Jullie zijn geweldig en onmisbaar.

Lieve Daniel, je houdt me keer op keer op het juiste pad.

Lieve Thijs, wat ben je een eerlijke en nieuwsgierige vriend zeg. Dank voor de mooie glossy die we samen hebben kunnen maken.

Lieve collega's van het Tergooi, Lucas, VUmc, GGZ-NHN, InGeest, DeJongePsychiater of andere werkzaamheden waar ik betrokken bij ben geweest. Dank voor jullie collegialiteit en inspiratie. Dankzij jullie reflecties en betrokkenheid werd ik iedere dag weer een iets betere psychiater.

Mijn lieve paranimfen Jur en Vinkers, dank voor de bijstand. Jullie onuitputtelijke wetenschappelijke en vriendschappelijke support in alle gekke projecten die we gezamenlijk ondernemen is keer op keer weer verfrissend en motiverend. Of het nu een nascholing is die we maken of organiseren, een brief of opinie die we schrijven, een artikel dat we produceren of een reis die we maken. Het blijven steeds weer vernieuwende avonturen waar ik geen genoeg van kan krijgen. Dank. Jullie maken me een creatiever wetenschapper en mens.

Beste medeauteurs, dank jullie wel! Voel je vooral niet gepasseerd want, als je dit leest, denk ik wel degelijk aan jou. Wetenschap is een teamsport en zonder jullie gedachten, talent, harde werk en teamgeest was ik verloren en was er helemaal niets van dit proefschrift terechtgekomen.

Geachte leden van de promotiecommissie, dank voor jullie interesse en tijd die jullie gestoken hebben in het beoordelen van mijn proefschrift. Dankzij mijn onderzoek weet ik dat jullie ontzettend druk zijn, veel stress en publicatiedruk ervaren en er altijd ontelbaar veel zaken op jullie liggen te wachten. Desalniettemin vonden jullie tijd om het proefschrift op inspirerende wijze te beoordelen. Ik heb veel aan jullie commentaren gehad.

Beste Yvo,
Als er iemand is die ik daadwerkelijk langdurig en met doortastendheid wil bedanken ben jij het. Zonder je steun en geloof in mijn ideeën in 2012 was ik aan de wolven overgeleverd. Jij leidde ze in goede banen, inspireerde me, gaf me vertrouwen, wist me keer op keer te wijzen op mijn onvolkomenheden, wist te relativeren, te passioneren en vooral met heel veel plezier wetenschap te bedrijven. Zonder dit plezier en jouw enthousiasme was mijn promotietraject geen succes geworden. Ik denk dat iedere promovendus die bij jou promoveert, heel erg blij mag zijn met zo'n betrokken professor en rolmodel. Ik hoop dat we in de toekomst nieuwe, out-of-the-box en inspirerende projecten kunnen opzetten waarbij we onze kennis samenvoegen om ook patiëntenzorg beter te maken.

Beste Lex,

Je was bereid om halverwege mijn promotietraject in te stappen en daar ben ik je erg dankbaar voor. Je hebt me in de afgelopen 2 jaar veel geleerd. Epidemiologische inzichten, methodologische standvastigheid, de juiste toon aanslaan in e-mails en in belangrijke vergaderingen. Je leerde me over leiderschap, wijsheid, ambitie, precisie, resistentie voor tegenslagen en bovenal over oprecht samenwerken. De meest inspirerende leiders hebben altijd oog voor het hele team en gunnen credits aan collega's met wie hij samenwerkt. Jij bent zo'n bijzondere leider. Daardoor voel je je als collega gewaardeerd en blijf je loyaal. Dat soort type leiders zijn er niet zoveel, vaak wordt leiderschap vertroebeld door narcisme of eigenbelang. Bij jou niet. Ik mag hopelijk nog lang met je samenwerken om van je te leren.

Voor nog meer mooie woorden van en over mijn promotoren verwijs ik graag naar 'the JOERI'. Dé wetenschapsglossy die ik eenmalig uitgeef naast dit proefschrift.

In de meeste medische proefschriften begint een dankwoord met het bedanken van de patiënten. Mijn 'patiënten' waren alle wetenschappers die hebben deelgenomen aan mijn onderzoeken. Zonder jullie heb ik dit onderzoek nooit kunnen uitvoeren. Zonder jullie waren er geen resultaten. Zonder jullie geduld en tijd waren de bladzijden in het proefschrift niet gevuld. Ik heb daarom even zitten rekenen. 437 hoogleraren hebben in het onderzoek in **hoofdstuk 2,3,en 4** meegedaan. De betreffende survey duurde gemiddeld 17 minuten. Dat zijn bijna 124 hoogleraar uren. In **hoofdstuk 5** hebben 100 hoogleraren, 100 postdocs en 120 PhD studenten meegedaan in een 18 minuten durende survey. Dat is 30 hoogleraar-uren en 66 overige wetenschappers-uren. **In hoofdstuk 6** hebben bijna 100 wetenschappers deelgenomen aan een 2 uur durende focus groep. Dat zijn 200 wetenschappers-uren. In **hoofdstuk 7 h**ebben bijna 400 psychiaters deelgenomen aan een 22 minuten durende vragenlijst. Dat is bijna 147 psychiater-uren. Daarna hebben in **hoofdstuk 8 en 9** 535 wetenschappers deelgenomen aan een 29 minuten durende vragenlijst. Dat zijn in totaal 259 wetenschappers-uren. Dan kom ik uit op een totaal van 825 wetenschappers-uren. Dat zijn bijna 104 werkdagen. Dat zijn bijna 21 volledige full-time werkweken. Ik ben heel erg dankbaar dat jullie op deze manier hebben kunnen bijdragen aan mijn proefschrift.

Tot slot, dank iedereen die ik vergeten ben te benoemen in het dankwoord. Er zit geen intentie achter, het is onbewust en dat ik u vergeten ben is mijn fout. Voel je vooral niet gepasseerd of gekrenkt. Het wordt mogelijk veroorzaakt door mijn selectief geheugen, door selectiebias, door promotiestress of in sporadische gevallen vanwege het feit dat u onvoldoende indruk heeft gemaakt om u in het dankwoord op te nemen.

Curriculum Vitae

Curriculum Vitae

Joeri Kees Tijdink was born on June 15th 1981 in Tiel. After a lighthearted childhood, he finished in 1999 his VWO at Lek & Linge School in Culemborg. In that same year he attended medical school at the University of Utrecht. In 2004, as a medical student, he worked for 12 months in San Juan de Dios Hospital in Guatemala DF, Guatemala, to complete a major part of his internal rotations. Shortly after graduating from medical school in November 2006, he started his clinical residencies in psychiatry at St. Lucas Andreas Hospital in Amsterdam under the supervision of Dr. Alex Korzec and Dr. Ton Vergouwen.

In 2012, in his last year of his clinical residencies, Joeri was intrigued by the academic enterprise. His deep fascination for academic subculture stretches from the belligerent and hostile aspects of the academic environment in University Medical Centers to the cultural psycho-socio-economic value of authorships for individual scientists. This has initially led to a collaboration with Prof. Y. Smulders and in a latter phase Prof. L. Bouter to study publication pressure, publication culture and research integrity. Under their supervision and dedication, Joeri managed to finish the doctorate in November 2015. During this period Joeri also worked as a clinical psychiatrist at Tergooi Hospitals and was responsible for the residency program in that same hospital. In 2015/2016 he worked a brief period at the Henry Rongomau Bennett Center in Hamilton, New Zealand as a consultant psychiatrist. Furthermore, he is the co-founder and editor of the Dutch platform for young psychiatrists (www.dejongepsychiater.nl, @jongepsychiater) that focusses on translation of psychiatric research into clinical practice and he directed the former and upcoming shows of Daniel Arends, professional comedian and performer.

To give a brief insight in the personality of Joeri Tijdink; his scores on the Dark Triad were (see chapter 9 for comparison): Machiavellianism 25 (moderate), Narcissism 25 (moderate) and Psychopathy 14 (low).

Joeri is living in Amsterdam, together with Roos van Grieken and their daughter Evy Tijdink.