Prevalence of responsible research practices among academics in the Netherlands

Gowri Gopalakrishna*, Jelte M. Wicherts, Gerko Vink, Ineke Stoop, Olmo R. van den Akker, Gerben ter Riet, Lex M. Bouter

Affiliations:

1. Gowri Gopalakrishna is at the Amsterdam University Medical Centers, Department of Epidemiology and Data Science, 1081 HV Amsterdam, The Netherlands.

2. Jelte M. Wicherts is at the Tilburg University, Department of Methodology and Statistics, 5000 LE Tilburg, The Netherlands.

3. Gerko Vink is at the Utrecht University, Department of Methodology and Statistics, 3584 CH Utrecht, The Netherlands.

4. Ineke Stoop is a retired Methodologist from The Netherlands Institute for Social Research, 2500 BD Den Haag, The Netherlands.

5. Olmo R. van den Akker is at the Tilburg University, Department of Methodology and Statistics, 5000 LE Tilburg, The Netherlands.

6. Gerben ter Riet is at the Amsterdam University of Applied Science, Center of Expertise Urban Vitality, Faculty of Health, 1097 DZ Amsterdam, The Netherlands.

7. Lex M. Bouter is at the Vrije Universiteit Amsterdam, Department of Philosophy, Faculty of Humanities and at the Amsterdam University Medical Centers, Department of Epidemiology and Data Science, 1081 HV Amsterdam, The Netherlands.

*Corresponding author. Email: g.gopalakrishna@amsterdamumc.nl

Summary Paragraph

Public trust in research is essential in order to meet important challenges of society. Trust in research is earned through transparent, responsible research practices. Traditionally, research integrity studies have focused on research misbehaviors and their explanations [1,2]. Over time attention has shifted from detecting and sanctioning research misconduct towards preventing questionable research practices [3]. So far little attention has been given to responsible research practices and how these can be fostered optimally. Especially in regards to open methods, open codes and open data [4,5]. We present results from a large survey [6] among academics in The Netherlands showing that prevalence of responsible practices differed substantially with almost all researchers declaring to avoid plagiarism in their work and only a minority pre-registering a research protocol prior to commencing data collection. Arts and humanities scholars as well as PhD candidates and junior researchers engaged less often in responsible research practices. Our findings suggest that publication pressure [7] affects responsible practices negatively, while mentoring, scientific norm subscription and funding pressure [8-10] may stimulate them. These results can help us to understand ways by which responsible research practices can be made more universal across disciplines and academic ranks so as to increase transparency and trustworthiness of research.

Introduction

Public trust in research is essential in order to meet important challenges of society. Trust in research is earned by being transparent as well as by performing research that is relevant, ethically sound and of high quality. Researchers and their research institutions should make research trustworthy by promoting responsible research practices (RRPs) and by discouraging questionable research practices (QRPs) and research misconduct [11]. To this end, solid, empirical knowledge on the adoption of RRPs and their underlying explanatory factors is essential.

There has been a clear rise in publications and efforts aimed at promoting research integrity in recent years [4,5,11-16], including pleas for the adoption and promotion of open science and other RRPs aimed at increasing the trustworthiness of research through increased transparency. In particular, open methods (e.g. preregistration of study protocols), open codes (for data analysis), open data (following the FAIR principles [17]) and open access (rendering publications available at no cost for users) play an important role [4].

A number of explanatory factors such as scientific norm subscription, fair distribution of resources, rewards and recognitions (i.e. organizational justice), perceived pressures researchers face (e.g. competition, work, publication and funding pressures), and support by mentors have been suggested to be important in fostering high quality research [18-20]. So far however, the body of research on research integrity has focused largely on how to minimize QRPs and not much on empirical evidence to foster RRPs. These studies have typically a narrow disciplinary scope covering few possible explanatory factors [1,9,10,18-22].

The National Survey on Research Integrity (NSRI) [6] was designed to take a balanced, research-wide approach to report on the prevalence of RRPs, QRPs and research misconduct in addition to exploring the potential explanatory factors associated with these behaviors in a

single survey. The NSRI targets the entire population of academic researchers in The Netherlands, across all disciplinary fields and academic ranks.

The objectives of the NSRI are:

1) to estimate prevalence of RRPs, QRPs and research misconduct and

2) to study the association between possible explanatory factors and RRPs, QRPs and research misconduct.

In this paper we focus on the prevalence of RRPs and the explanatory factors that may help or hinder responsible conduct of research. Elsewhere we report on QRPs, research misconduct and their associative explanatory factors [23].

Results

Descriptive analyses

63,778 emails were sent out (Figure 1) and 9529 eligible respondents started the survey. Of these, 2716 stopped the survey prematurely and 6813 completed the survey. The response could be reliably calculated for the eight supporting institutions only (Supplementary Figure 1a) and was 21.2%.

Supplementary Table 1a gives a breakdown of all respondents stratified by background characteristics. Male and female respondents are fairly equally split among the respondents. For the natural and engineering sciences, women account for 24.9% of respondents. In the highest academic rank of associate and full professors, women make up less than 30% of respondents (Supplementary Table 1a). Nearly 90% of all respondents are engaged in empirical research and about half (48%) come from the eight supporting institutions. Respondents from supporting and non-supporting institutions are fairly evenly distributed

across disciplinary fields and academic ranks except for the natural and engineering sciences where less than one in four (23.5%) come from supporting institutions.

PhD candidates and junior researchers have the lowest scale score for work pressure (3.9) compared to the other ranks (Supplementary Table 1b). Postdocs and assistant professors reported the highest scale scores for publication pressure (4.2), funding pressure (5.2), and competitiveness (3.7) and the lowest organizational justice (4.1) scores compared to the other ranks (Supplementary Table 1b).

Respondents from the arts and humanities have higher scale scores for work pressure (4.8), publication pressure (4.1) and competitiveness (3.8) and the lowest scale scores for mentoring and organizational justice (3.5 and 3.9, respectively) (Supplementary Table 1b). The scientific norms scale scores are similar across all disciplines and academic ranks. The scores on the peer norms scale are consistently lower than the scientific norms scores across disciplines and ranks.

Prevalence of RRPs

The five most prevalent RRPs (i.e. with a Likert scale score of 5, 6 or 7) have a prevalence range of 86.4% to 99% (Table 1, Supplementary Figure 2). Fair ordering of authorships (RRP 3) and preregistration of study protocols (RRP 6) show the largest percentage difference between the life and medical sciences and the arts and humanities (RRP 3: 75.7 vs 91.6% and RRP 6: 50.8% vs 30.2%). PhD candidates and junior researchers (74.2%) report the lowest prevalence for RRP3 on fair allocation of authorships compared to associate and full professors (90.9%).

Supplementary Table 4 shows the discipline-specific and academic rank-specific prevalence of "not applicable" (NA) answers on the 11 RRPs. Arts and humanities scholars report the

highest prevalence of NA for 9 out of the 11 RRPs. Similarly across ranks, PhD candidates and junior researchers display the highest prevalence of NAs on 9 out of the 11 RRPs.

The four open science practices have an overall prevalence ranging from 42.8% to 75%: (i) following the FAIR principles (RRP 4: 75%); (ii) Publishing open access (RRP 8: 72.6%); (iii) Providing underlying data, computer codes, or syntaxes (RRP 10: 47.2%) and (iv) Preregistration of study protocols (RRP 6: 42.8%) (Table 1).

Surprisingly, the arts and humanities scholars have the highest prevalence for RRP 4 on following FAIR principles (84.6%). However a closer look at RRP 4, reveals that this discipline also has the highest percentage of NA for RRP 4 (27.5%) (Supplementary Table 4). Life and medical sciences have the highest prevalence (50.8%) and the arts and humanities the lowest (30.2%) for preregistration of study protocols (RRP 6) where nearly 70% (67.8%) of the arts and humanities scholars rated RRP 6 as not applicable (Supplementary Table 4). Arts and humanities scholars have the lowest prevalence (59.1%) and the life and medical sciences the highest (75.1%) for publishing open access (RRP 8) (Table 1).

Regression analyses

Table 2a shows the results of the linear regression analysis for the five background characteristics while Table 2b shows the results for the explanatory factor scales.

Table 2a shows that the arts and humanities scholars have a significantly lower overall RRP mean score (-0.51; 95% CI -0.59, -0.42). Similarly, doing non-empirical research is associated with a significantly lower overall RRP mean score (-0.49; 95% CI-0.57, -0.42). Females have a significantly lower RRP mean score than males (-0.07; 95% CI -0.12, -0.02). Being a PhD candidate or junior researcher is associated with a significantly lower overall RRP mean

(-0.31; 95% CI -0.37, -0.25).

One standard deviation increase on the publication pressure scale is associated with a significant decrease in overall RRP mean score (-0.05; 95% CI -0.08, -0.02) (Table 2b). An increase of one standard deviation in the following five explanatory factor scales is associated with higher overall RRP mean, namely: (i) mentoring (0.15; 95% CI 0.12, 0.17); (ii) funding pressure (0.13; 95% CI 0.10, 0.17); (iii) scientific norms subscription (0.13; 95% CI 0.10, 0.15); (iv) likelihood of QRP detection by collaborators (0.06; 95% CI 0.03, 0.09); and (v) work pressure (0.03; 95% CI 0.00, 0.06).

The pre-specified subgroups analyses did not show statistically significant differences in QRP mean score.

Discussion

Summary of main findings

We found that overall RRP prevalence ranged from 42.8% to 99% with open science practices at the lower end (42.8% to 75%). The arts and humanities scholars have the lowest prevalence of preregistration of study protocols and open access publication. This disciplinary field also has the highest prevalence of NAs (9 out of the 11 RRPs) as do the PhD candidates and junior researchers. Arts and humanities scholars as well as PhD candidates and junior researchers are associated with a significantly lower overall RRP mean score as is doing non-empirical research and being female in gender.

Publication pressure is associated with lower overall RRP mean score while mentoring, funding pressure, scientific norm subscription, likelihood of QRP detection by collaborators and work pressure are associated with higher RRP mean scores.

Explanatory factors that promote or hinder RRP

The results of our regression analysis suggest that publication pressure might lower RRPs, although the effect is modest. This finding [23] complements what we found for QRPs where publication pressure was associated with a higher odds of engaging frequently in at least one QRP. These results suggest that lowering publication pressure may be important for fostering research integrity.

Noteworthy are our findings regarding scientific norm subscription and peer norms [18,20]. These scales have been previously validated and used in a study among 3600 researchers of different disciplines in the United States of America [8,20]. In that study, respondents reported higher scientific norm subscriptions when asked about the norms a researcher should espouse but they perceived the actual adherence to these norms by their peers to be lower. Our results corroborate these findings [20].

Previous authors have made calls to institutional leaders and department heads to pay increased attention to scientific norms subscription within their research cultures [20,24]. Our regression analysis findings reinforce these calls to revive subscription to the Mertonian scientific norms [8].

Mentoring is associated with a higher overall RRP mean score. Interestingly, a lack of proper supervision and mentoring of junior co-workers is the third most prevalent QRPs respondents reported in our survey [23]. This finding was also reported in another recent survey among researchers in Amsterdam [25]. These findings suggest that increased efforts to improve mentoring and supervision may be warranted within research institutions.

In our QRP analysis of the NSRI survey results, likelihood of detection by reviewers was significantly associated with less misconduct, suggesting that reviewers, more than

collaborators, are important in QRP detection [23]. However, for RRPs, the reverse seems to be true: collaborators more so than reviewers may be important for fostering RRPs.

To our surprise we find that work pressure and funding pressure both have a small but significant association with higher RRP mean scores. One plausible explanation may be that adhering to RRPs requires a slower, more meticulous approach to performing research responsibly.

Open science practices

We found that scholars from the arts and humanities, as well as PhD candidates and junior researchers reported RRPs more often as "not applicable". We are unable to differentiate whether this is because these open science RRPs are truly not applicable or if these practices are simply not yet recognized as standard responsible practices in this discipline and rank. While it can be argued that not all open science practices, particularly those relating to the sharing of data and codes, are relevant for the non-empirical disciplines such as the arts and humanities, practices like preregistration of study protocols, publishing open access and making sources, theories and hypotheses explicit and accessible, seem relevant for most types of research, empirical or not.

Areas of focus: background characteristics

Arts and humanities scholars reported the highest work pressure, publication pressure and competitiveness, and the lowest organizational justice and mentoring support. While our sample size for this disciplinary field is relatively small (n = 636), the finding of lower organizational justice in this discipline is consistent with a recent study [26]. Our regression analysis shows that arts and humanities scholars have significantly lower overall RRP mean scores and also the highest prevalence of not applicables for 9 out of the 11 RRPs. Research integrity efforts have largely focused on the biomedical and social behavioural sciences [27].

However, these results point to a need to better understand responsible research practices that may be disciplinary field-specific.

We found that PhD candidates and junior researchers have the lowest prevalence across all RRPs and are associated with the lowest overall RRP mean score. A recent Dutch survey of academics as well as our own survey point to inadequate mentoring and supervision of junior co-workers as a prevalent QRP [23,28]. This seems to underline a clear message: mentoring and supervision is a key aspect in the training of PhD candidates and junior researchers as it can help foster RRPs.

Women have a (small, yet statistically significant) lower overall RRP mean score. While it has been previously reported that men engage in research misbehavior more than women [23,25,29], our finding of lower RRP engagement for women has not been reported earlier and is a findings we intend to explore in the qualitative discussions planned in the next phase of our project.

Strengths and Limitations

The email addresses of researchers affiliated to non-NSRI-supporting institutions were web-scraped from open sources. Therefore, we are unable to credibly verify if the scraped email addresses matched our eligibility criteria for NSRI participation. Hence, we calculated the response based only on the eight supporting institutions. The 21.1% response is within the range of similar research integrity surveys [2,26]. Given this response, one may wonder how representative the NSRI sample is of the target population i.e. all academic researchers in the Netherlands. Unfortunately, there are no reliable numbers at the national level that match our study's eligibility criteria. Therefore, we cannot assess our sample's representativeness even for the five background characteristics. Nevertheless, we believe our results to be valid as our main findings align well with the findings of other research integrity surveys [10,20,24,26,30].

A limitation of our analysis concerns recoding NA answers into “never” for the multiple linear regressions since there is a difference between not committing a behaviour because it is truly not applicable and intentionally refraining from doing so. Our analyses may therefore underestimate the occurrence of true intentional RRPs.

The NSRI is the largest research integrity survey in academia to-date that has looked at both prevalence of RRPs as well as the largest range of explanatory factors in one single study across all disciplinary fields and academic ranks.

References

1 Xie, Y., Wang, K. & Kong, Y. Prevalence of research misconduct and questionable research practices: a systematic review and meta-analysis. *Science and Engineering Ethics* **27**, 41 (2021).

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

3 Bouter, L. M., Tijdink, J. K., Axelsen, N., Martinson, B. C. & ter Riet, G. Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. *Research Integrity and Peer Review* **1**, 17 (2016).

4 Moher, D. *et al.* The Hong Kong Principles for assessing researchers: Fostering research integrity. *PLoS Biology* **18**, e3000737 (2020).

5 Mejlgaard, N. *et al.* Research integrity: nine ways to move from talk to walk. *Nature* **586**, 358-360 (2020).

6 National Survey on Research Integrity. <www.nsri2020.nl Accessed Date: 13 July 2021>

7 Haven, T. L., de Goede, M. E. E., Tijdink, J. K. & Oort, F. J. Personally perceived publication pressure: revising the Publication Pressure Questionnaire (PPQ) by using work stress models. *Research Integrity and Peer Review* **4**, 7 (2019).

8 Anderson, M. S., Ronning, E. A., De Vries, R. & Martinson, B. C. Extending the Mertonian norms: scientists' subscription to norms of research. *The Journal of Higher Education* **81**, 366-393 (2010).

9 Anderson, M. S., Ronning, E. A., De Vries, R. & Martinson, B. C. The perverse effects of competition on scientists' work and relationships. *Science and Engineering Ethics* **13**, 637-461 (2007).

10 Anderson, M. S. *et al.* What do mentoring and training in the responsible conduct of research have to do with scientists' misbehavior? Findings from a national survey of NIH-funded scientists. *Academic Medicine* **82**, 853-860 (2007).

11 National Academy of Sciences, National Academy of Engineering, Institute of Medicine, Committee on Science and Engineering and Public Policy & Panel on Scientific Responsibility and the Conduct of Research. *Responsible science - Ensuring the integrity of the research process: volume 1*. Vol. 1 (1992).

12 Steneck, N., Mayer, T. & Anderson, M. S. Singapore statement on research integrity. (Singapore, 2010).

13 Shaw, D. M. & Thomas, C. Ten simple rules for protecting research integrity. *PLoS Computational Biology* **11**, e1004388 (2015).

14 Hiney, M. Briefing paper on research integrity: what it means, why it is important and how we might protect it. (Science Europe, 2015). <https://www.scienceeurope.org/our-resources/briefing-paper-on-research-integrity-what-it-means-why-it-is-important-and-how-we-might-protect-it Accessed Date: 13 July 2021>

15 Committee on Responsible Science, Committee on Science Engineering Medicine and Public Policy, Policy and Global Affairs & National Academies of Sciences Engineering and Medicine. *Fostering integrity in research*. (The National Academies Press, 2017).

16 ALLEA - All European Academies. The european code of conduct for research integrity (revised edition). (Berlin, 2017).

17 Wilkinson, M. D. *et al.* The FAIR guiding principles for scientific data management and stewardship. *Scientific Data* **3** (2016).

18 Merton, R. K. *The sociology of science: theoretical and empirical investigations*. (University of Chicago Press, 1973).

19 KNAW, NFU, NOW, TO2-federatie, Vereniging Hogescholen, VSNU. Netherlands code of conduct for research integrity. (The Netherlands, 2018).

20 Anderson, M. S., Martinson, B. C. & De Vries, R. Normative dissonance in science: results from a national survey of U.S. scientists. *Journal of Empirical Research on Human Research Ethics* **2**, 3-14 (2007).

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

22 Bonn, N. A. & Pinxten, W. Advancing science or advancing careers? Researchers' opinions on success indicators. *PLoS One* **16**, e0243664 (2021).

23 Gopalakrishna, G. *et al.* Prevalence of questionable research practices, research misconduct and their potential explanatory factors: a survey among academic researchers in The Netherlands. (2021). < https://osf.io/preprints/metaarxiv/vk9yt Accessed Date: 13 July 2021>

24 Martinson, B. C., Crain, A. L., De Vries, R. & Anderson, M. S. The importance of organizational justice in ensuring research integrity. *Journal of Empirical Research on Human Research Ethics* **5**, 67-83 (2010).

25 Fang, F. C., Bennett, J. W. & Casadevall, A. Males are overrepresented among life science researchers committing scientific misconduct. *American Society for Microbiology* **4**, e0064012 (2013).

26 Haven, T. L., Tijdink, J. K., Martinson, B. C. & Bouter, L. M. Perceptions of research integrity climate differ between academic ranks and disciplinary fields: Results from a survey among academic researchers in Amsterdam. *PLoS One* **14**, e0210599 (2019).

27 Bonn, N. A. & Pinxten, W. A decade of empirical research on research integrity: what have we (not) looked at? *Journal of Empirical Research on Human Research Ethics* **14**, 338-352 (2019).

28 Haven, T. L. *et al.* Researchers' perceptions of research misbehaviours: a mixed methods study among academic researchers in Amsterdam. *Research Integrity and Peer Review* **4**, 25 (2019).

29 Charness, G. & Gneezy, U. Strong evidence for gender differences in risk taking. *Journal of Economic Behavior & Organization* **83**, 50-58 (2012).

30 van Buuren, S. & Groothuis-Oudshoorn, K. mice: multivariate Imputation by chained equations in R. *Journal of Statistical Software* **45**, 1-67 (2011).

31 National Survey on Research Integrity on Open Science Framework (OSF). <https://osf.io/ehx7q/ Accessed Date: 13 July 2021>

32 Kantar Public. <http://www.nipo.nl/ Accessed Date: 13 July 2021 >

33 ESOMAR. Kantar signs up to ICC/ESOMAR international code globally - inks new membership deal focused on employee development. (2020).

34 Navarro, M. L. A., Mas, M. B. & Jiménez, A. M. L. Working conditions, burnout and stress symptoms in university professors: Validating a structural model of the mediating effect of perceived personal competence. *The Spanish Journal of Psychology* **13**, 284-296 (2010).

35 Miller, K., Willson, S., Chepp, V. & Padilla, J. L. Cognitive interviewing methodology. (John Wiley & Sons, Inc, Hoboken, New Jersey, 2014).

36 Cruyff, M. J. L. F., Van den Hout, A. D. L. & Van der Heijden, P. G. M. The analysis of randomized-response sum score variables. *Journal of the Royal Statistical Society. Series B: Statistical Methodology* **70**, 21-30 (2008).

37 Rubin, D. B. Statistical matching using file concatenation with adjusted weights and multiple imputations. *Journal of Business & Economic Statistics* **4**, 87-94 (1986).

38 Little, R. J. A. Missing-data adjustments in large surveys (with discussion). *Journal of Business Economics and Statistics* **6**, 287–296 (1988).

39 Rubin, D. B. *Multiple imputation for nonresponse in surveys*. 76 (John Wiley & Sons, 1987).

40 Mitroff, I. I. *The subjective side of science: a philosophical inquiry into the psychology of the Apollo Moon Scientists*. (Elsevier Scientific Publishing Company, 1974).

Figure 1: Flow chart of the survey

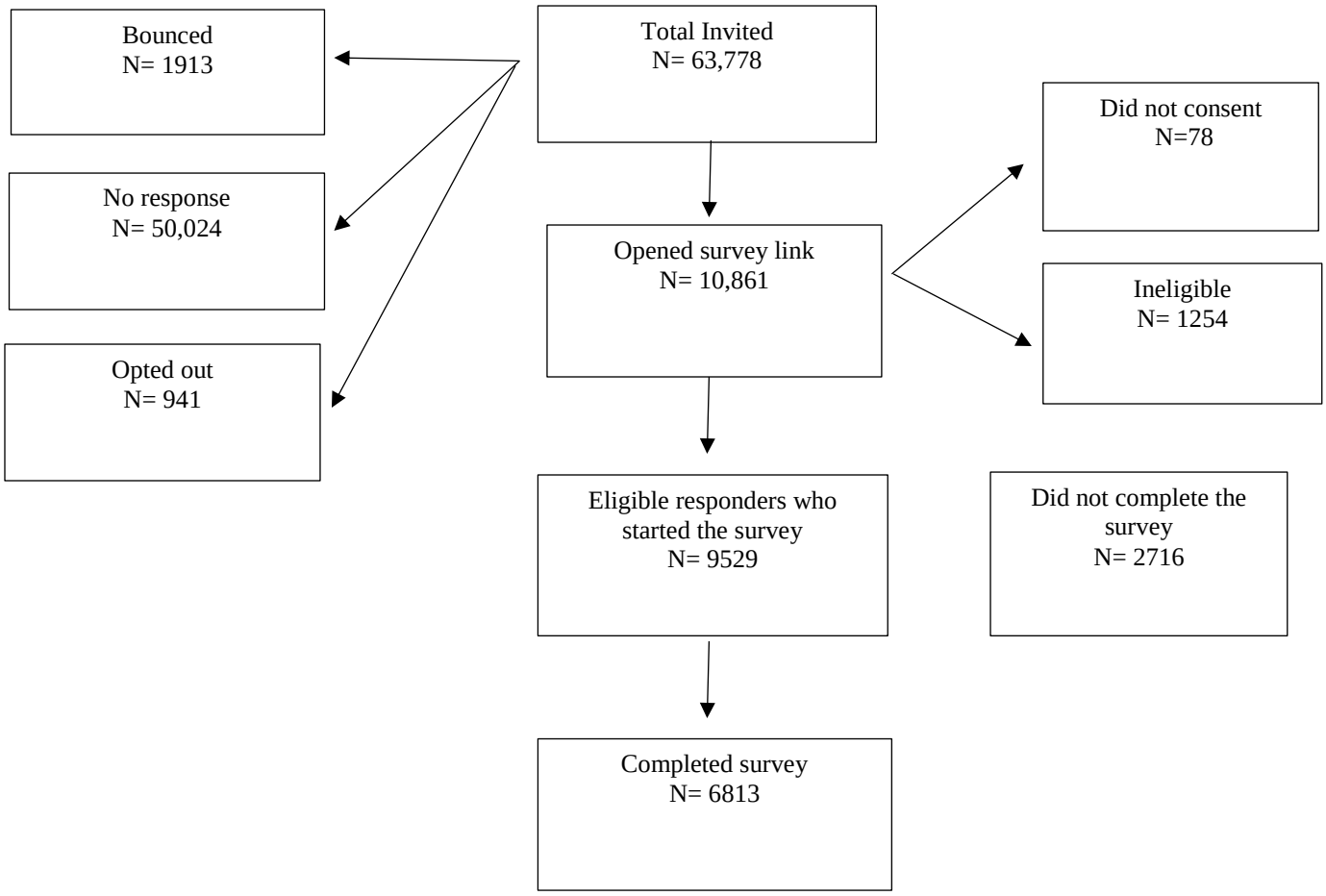

Table 1: Estimated prevalence[+] (95% confidence intervals) of the 11 RRPs stratified by disciplinary field and academic rank[^]

		Disciplinary field				Academic rank			
RRP	**Description (In the last three years..)**	**Life and medical sciences**	**Social and behavioral sciences**	**Natural and engineering sciences**	**Arts and humanities**	**PhD candidates and junior researchers**	**Postdocs and assistant professors**	**Associate and full professors**	**Overall**
RRP1	I disclosed who funded my studies and all my relevant financial and non-financial interests in my publications	98.6 (98,99)	96.2 (95.1,97)	94 (92.6,95.2)	93.2 (90.3,95.3)	94 (92.6,95.1)	97.3 (96.6,97.9)	97.5 (96.7,98.2)	96.5 (96,97)
RRP2	I took steps to correct errors in my published work whenever I and/or peers provided valid reasons for such a correction	88.9 (87.1,90.5)	83.4 (80.7,85.8)	85.5 (82.9,87.8)	86.5 (82,90)	87.9 (85.5,89.9)	84.5 (82.5,86.4)	87.7 (85.6,89.6)	86.4 (85.2,87.6)
RRP3	The allocation and ordering of authorships in my publications, were fair and in line with the standards of my discipline	75.7 (74,77.3)	84.1 (82.4,85.8)	86.6 (84.7,88.3)	91.6 (88.7,93.8)	74.2 (72.1,76.3)	79.6 (78,81.1)	90.9 (89.5,92.1)	81.8 (80.8,82.7)
RRP4	I contributed, where appropriate, to making my research data findable, accessible, interoperable and reusable in accordance with the FAIR principles	74.8 (73.1,76.5)	70.7 (68.4,72.8)	77.5 (75.1,79.7)	84.6 (80.9,87.7)	75.2 (73,77.4)	73.6 (71.8,75.3)	76.6 (74.6,78.4)	75 (73.9,76.1)
RRP5	I kept a comprehensive record of my research decisions throughout my studies.	57.2 (55.3,59.2)	56.5 (54.2,58.8)	54 (51.2,56.7)	57.1 (52.5,61.6)	62.2 (59.9,64.4)	56.4 (54.4,58.3)	50.4 (48.1,52.7)	56.3 (55.1,57.6)
RRP6	I pre-registered my study protocols in line with open science practices	50.8 (48.5,53.1)	38.9 (36.3,41.6)	31.9 (28.4,35.5)	30.2 (24.1,37.1)	44.3 (41.4,47.3)	40 (37.7,42.4)	45.2 (42.5,47.9)	42.8 (41.3,44.3)
RRP7	I managed my research data carefully by storing both the raw and processed versions for a period appropriate to my discipline and methodology used	90.9 (89.7,91.9)	88.8 (87.2,90.2)	84.5 (82.4,86.5)	82.8 (78.7,86.3)	90.8 (89.3,92)	87.9 (86.5,89.1)	86.7 (85.1,88.3)	88.4 (87.6,89.2)
RRP8	My research was published under open access conditions	75.1 (73.3,76.8)	72.7 (70.6,74.8)	73.7 (71.2,76)	59.1 (54.9,63.2)	73.8 (71.4,76.1)	72 (70.3,73.7)	72.6 (70.6,74.5)	72.6 (71.5,73.7)
RRP9	When making use of other people's ideas, procedures, results and text in my publications, I cited the source accurately in accordance with the standards of my discipline	98.8 (98.3,99.2)	99.3 (98.8,99.6)	98.9 (98.1,99.3)	99.4 (98.2,99.8)	98.8 (98.2,99.2)	98.8 (98.3,99.1)	99.5 (99.1,99.8)	99 (98.7,99.2)
RRP10	I fully disclosed and made accessible on open science platforms my underlying data, computer codes, or syntaxes used in my research	47.4 (45.2,49.5)	41.4 (38.8,44.1)	52.7 (49.8,55.6)	53.4 (46.3,60.3)	42.4 (39.6,45.2)	47.1 (44.9,49.2)	51 (48.6,53.5)	47.2 (45.8,48.6)
RRP11	Before releasing results of my research, I meticulously checked my work to avoid errors and biases	94.3 (93.4,95.2)	94.8 (93.6,95.7)	93.6 (92.2,94.8)	94.2 (92,95.9)	94.3 (93.1,95.3)	94.4 (93.4,95.2)	94.2 (93,95.1)	94.3 (93.7,94.8)

5

[+] Prevalence is based on the RRP at issue having a Likert score of 5, 6 or 7 among respondents that deemed the RRP at issue applicable; [^]All figures in this table are percentages and refer to the last 3 years.

Table 2a: Linear regression coefficients (95% confidence interval) of overall RRP mean score^ stratified by background characteristics

		Overall RRP mean score
		Linear regression model†† Mean difference from reference category (95% CI)
Disciplinary field	**Social and behavorial sciences**	**-0.15 (-0.20, -0.10)**
Reference category: Life and medical sciences	Natural and engineering sciences	-0.03 (-0.09, 0.04)
	Arts and humanities	**-0.51 (-0.59, -0.42)**
Academic rank	**PhD candidates and junior researchers**	**-0.31 (-0.37, -0.25)**
Reference category: Postdocs and assistant professors	**Associate and full professors**	**0.08 (0.03, 0.14)**
Gender	**Female**	**-0.07 (-0.12, -0.02)**
Reference category: Male	Undisclosed	0.07 (-0.10, 0.24)
Engaged in empirical research	**No**	**-0.49 (-0.57, -0.42)**
Reference category: Yes		
Institutional Support	**Yes**	**-0.06 (-0.1, -0.01)**
Reference category: No		

^Overall RRP mean score was computed as the average score on the 11 RRPs with the not applicable scores recoded to 1 (i.e. never); ††Model containing the five background variables and all 10 explanatory factor scales; Bold figures are statistically significant..

Table 2b: Linear regression coefficients (95% confidence intervals) of overall RRP mean score[^] by explanatory factor scales

	Overall RRP mean score[^]
	Linear regression model Change in mean score per standard deviation increase (95 % CI)[††]
Work pressure	**0.03 (0.00, 0.06)**
Publication pressure	**-0.05 (-0.08, -0.02)**
Funding pressure	**0.13 (0.10, 0.17)**
Mentoring *	**0.15 (0.12, 0.17)**
Competitiveness	0.02 (-0.01, 0.05)
Scientific norm	**0.13 (0.10, 0.15)**
Peer norms	0.01 (-0.02, 0.04)
Organizational justice **	0.03 (-0.01, 0.06)
Likelihood of detection (collaborators)	**0.06 (0.03, 0.09)**
Likelihood of detection (reviewers)	0.00 (-0.03, 0.02)

[^] *Overall RRP mean score was computed as the average score on the 11 RRPs with the not applicable scores recoded to 1 (i.e. never);* [††]*Model containing the five background variables (see Table 2a) and all 10 explanatory factor scales; *Two scales (Responsible Mentoring and Survival Mentoring) were merged due to high correlation ; **Two subscales (Distributional and Procedural Organizational Justice) were merged due to high correlation; Supplementary Table 3 shows the correlation of all the explanatory factor scales; Bold figures are statistically significant.*

Methods

Ethics approval

The Ethics Review Board of the School of Social and Behavioral Sciences of Tilburg University approved this study (Approval Number: RP274). The Dutch Medical Research Involving Human Subjects Act (WMO) was deemed not applicable to this study by the Institutional Review Board of the Amsterdam University Medical Centers (Reference Number: 2020.286).

The full NSRI study protocol, ethics approvals, complete data analysis plan and final dataset can be found at [31]. Below we summarize the salient study features.

Study Design

The NSRI is a cross-sectional study using a web-based anonymized questionnaire. All academic researchers working at or affiliated to at least one of 15 universities or 7 UMCs in The Netherlands were invited by email to participate. To be eligible, researchers had, on average, to do at least 8 hours of research-related activities weekly, belong to life and medical sciences, social and behavioural sciences, natural and engineering sciences, or the arts and humanities and had to be a PhD candidate or junior researcher, postdoctoral researcher or assistant professor, or associate or full professor.

The survey was conducted by a trusted third party, Kantar Public [32], which is an international market research company that adheres to the ICC/ESOMAR International Code of standards [33]. Kantar Public's sole responsibility was to send the survey invitations and reminders by email to our target population and send the anonymized dataset at the end of the data collection period to the research team.

Universities and UMCs that supported NSRI supplied Kantar Public with the email addresses of their eligible researchers. Email addresses for the other institutes were obtained through publicly available sources, such as university websites and PubMed.

Researchers' informed consent was sought through a first email invitation which contained the survey link, an explanation of NSRI's purpose and its identity protection measures. Consenting invitees could immediately participate. NSRI was open for data collection for seven weeks, during which three reminder emails were sent to non-responders, at a one to two week interval period. Only after the full data analysis plan had been finalized and preregistered on the Open Science Framework [31] did Kantar Public send us the anonymized dataset containing individual responses.

Survey Instrument

NSRI comprises of four components: 11 QRPs, 11 RRPs, two FFs and 12 explanatory factor scales (75 questions). The survey starts with a number of background questions to assess eligibility of respondents. These include questions on one's weekly average duration of research-related work, one's dominant field of research, academic rank, gender and if one is doing empirical research or not [31].

All respondents, regardless of their disciplinary field or academic rank were presented with the same set of RRPs, QRPs and research misconduct questions. These questions referred to the last three years in order to minimize recall bias. The 11 RRPs were adapted from the Dutch Code of Conduct for Research Integrity 2018 [19] and a survey among participants of the World Conferences on Research Integrity [3]. GG created the initial formulations of the RRPs which cover study design, data collection, reporting, open science practices, conflicts of interest and collaboration. These 11 RRP formulations were reviewed and agreed upon in two rounds: first

within the NSRI core research team, and subsequently by an external group of multidisciplinary experts who formed the NSRI Steering Committee [6]. All 11 RRPs had a 7-point Likert scale ranging from 1 = never to 7 = always, in addition to a "not applicable" (NA) answer option.

The explanatory factors scales were based on psychometrically tested scales in the research integrity literature and focused on action-ability. Twelve were selected: scientific norms, peer norms, perceived work pressure, publication pressure, pressure due to dependence on funding, mentoring (responsible and survival), competitiveness of the research field, organizational justice (distributional and procedural), and likelihood of QRP detection by collaborators and reviewers [3,7,8,18-20,24,34].

Some of the scales were incorporated into the NSRI questionnaire verbatim, others were adapted for our population or newly created (see Supplementary Table 5). The scales on scientific norms, peer norms, competitiveness, organizational justice, and perceived likelihood of QRP detection were piloted.

We used "missingness by design" to minimize survey completion time. Thus, each invitee received one of three random subsets of 50 explanatory factor items from the full set of 75 (see Supplementary Table 5). All explanatory factor items had 7-point Likert scales. In addition, the two perceived likelihood of QRP detection scales, the procedural organizational justice scale and the funding pressure scale had a NA answer option. There was no item non-response as respondents had to either complete the survey or withdraw. We pre-tested the NSRI questionnaire's comprehensibility in cognitive interviews [35] with 18 academics from different ranks and disciplines.

Statistical analysis

We report on RRPs both in terms of prevalence and overall RRP mean. Prevalence we operationalized as the proportion of participants that scored 5, 6 or 7 among the participants that deemed the RRP at issue applicable. Means scores of individual RRPs only consider respondents that deemed the RRP at issue applicable. In the multiple linear regression analysis overall RRP mean was computed as the average score on the 11 RRPs, with the not-applicable scores recoded to 1 (i.e., "never"). Supplementary figures 2a to 2e show the distribution of responses, including the "not-applicable" category for the 11 RRPs. The associations of the overall RRP mean with the five background characteristics (Supplementary Table 1a) and the explanatory factor scales were investigated with multiple linear regression [36].

For the multivariate analyses of the explanatory factor scales, we used z-scores computed as the first principal component of the corresponding items [30]. Missing explanatory factor item scores due to 'not applicable' answers were replaced by the mean z-score of the other items of the same scale. Multiple imputation with mice in R [30] (version 4.0.3) was employed to deal with the missingness by design. Fifty complete data sets were generated by imputing the missing values using predictive mean matching [37,38]. The linear regression models were fitted to each of the 50 data sets, and the results combined into a single inference. To incorporate uncertainty due to the nonresponse, the inferences were combined according to Rubin's Rules [39]. All models contain all explanatory scales and the five background characteristics. The full statistical analysis plan, and analysis codes were preregistered on the Open Science Framework [31] including the following pre-specified subgroup analyses: field by rank, publication pressure by rank, funding pressure by rank, competition by disciplinary field, and detection (by reviewers or by collaborators) by disciplinary field.

Identity protection

Respondents' identity was protected in accordance to the European General Data Protection Regulations (GDPR) and corresponding legislation in The Netherlands. In addition, we had Kantar Public conduct the survey to ensure that the email addresses of respondents were never handled by the research team. Kantar Public did not store respondents' URLs and IP addresses. Only a fully anonymized dataset was sent to the research team upon closure of data collection and preregistration of the statistical analysis plan. Finally, we conducted analyses at aggregate levels only (i.e., across disciplinary fields, gender, academic ranks, whether respondents conducted empirical research, and whether they came from NSRI supporting institutions).

Data and materials availability

All data are available in the main text or the supplementary materials and online at https://osf.io/ehx7q/.

Methods References

3, 6-8, 18-20, 24, 30-32, 34-39

Acknowledgments

The authors wish to thank the NSRI Steering Committee members (Guy Widdershoven, Herman Paul, Ineke Stoop, Joeri Tijdink, Sonja Zuijdgeest, Corrette Ploem) for their support. In addition, we wish to thank Sara Behrad, Frank Gerritse, Coosje Veldkamp, Brian Martinson and Melissa Anderson for their contributions.

Author contributions

Conceptualization: GG, GtR, IS, JMW, LMB

Methodology: GG, GtR, GV, IS, JMW, LMB

Investigation: GG, JMW, OvdA

Visualization: GG, GV, LMB

Funding acquisition: GtR, LMB

Project administration: GG, LMB

Supervision: GG, GtR, LMB

Writing – original draft: GG, GtR, LMB

Writing – review & editing: GG, GtR, GV, IS, JMW, LMB, OvdA

Competing interests

Authors declare that they have no competing interests.

Funding

1- Netherlands Organisation for Health Research and Development (ZonMw) 20-22600-98-401

2- Consolidator Grant 726361 (IMPROVE) from the European Research Council (ERC, https://erc.europa.eu), awarded to J.M. Wicherts