Do research misbehaviours differ between disciplinary fields?¹

A mixed methods study among academic researchers in Amsterdam

Authors: Tamarinde Haven¹, Joeri Tjldink¹², H.Roeline Pasman³, Guy Widdershoven², Gerben ter Riet⁴⁵, Lex Bouter¹⁶

1. Department of philosophy, Vrije Universiteit Amsterdam

2. Department of medical humanities, Amsterdam UMC, location VUmc

3. Department of Public and Occupational Health, Amsterdam UMC, location VUmc

4. Department of research and education - ACHIEVE, Amsterdam University of Applied Sciences, Amsterdam, The Netherlands

5. Department of cardiology, Amsterdam UMC, location Meibergdreef, Amsterdam, The Netherlands

6. Department of epidemiology and biostatistics, Amsterdam UMC, location VUmc

¹ This work will be presented at the 6th World Conference on Research Integrity.

Abstract

Background

There is increasing evidence that research misbehaviour is common, especially the minor forms. Previous studies on research misbehaviour primarily focused on biomedical and social sciences, and evidence in the natural sciences and humanities is scarce. We investigated whether the importance of particular research misbehaviours, as perceived by researchers in their respective disciplines, varies across disciplinary fields.

Methods

We used a mixed methods explanatory quantitative first design. First, survey participants from 4 disciplinary fields rated perceived frequency and impact of research misbehaviours from a list of 60. We then combined these into a top 5 ranking of most important research misbehaviours, stratified by disciplinary field. Second, in focus group interviews participants from each academic rank and disciplinary field were asked to reflect on the most important research misbehaviours for their disciplinary field. We used participative ranking methodology pressing participants to obtain consensus on which research misbehaviours are most important.

Results

The survey response rate was 15%. Between 3 and 8 researchers participated per focus group. Insufficient supervision consistently ranked highest in the survey regardless of disciplinary field and the focus groups confirmed this. Important themes in the focus groups were insufficient supervision, sloppy science, and sloppy peer review. Biomedical researchers and social science researchers were primarily concerned with sloppy science and insufficient supervision. Natural sciences and humanities researchers discussed sloppy reviewing and theft of ideas by reviewers, a form of plagiarism. Focus group participants further provided examples of particular research misbehaviours they were confronted with and how these impacted their work as a researcher.

45 **Conclusion**

46 Academic researchers considered insufficient supervision and a variety of sloppy science
47 examples the most important research misbehaviours. Research misbehaviour that involved
48 the stealing of ideas and misconduct in the review process, which were absent in the initial
49 list of 60, were identified as additional research misbehaviours. Future research should
50 investigate the relevance and importance of these added research misbehaviours.

51

52 **Keywords:** *Research misbehaviour, research integrity, disciplinary fields, academic ranks,*
53 *research misconduct, survey, focus groups*

Background

Most researchers believe to possess a considerable degree of integrity, honesty and objectivity (1–3). Notwithstanding this, there is increasing evidence that researchers misbehave quite frequently in their professional work (4–6). Besides on the classical ‘deadly sins of science’ (falsification, fabrication and plagiarism (FFP)) there is little evidence on what are perceived to be the most important research misbehaviours (7,8). It’s becoming increasingly clear that research misbehaviours that may seem minor compared to FFP could be major in their aggregate impact since they occur much more frequently than the ‘deadly sins’ (9–12).

A meta-analysis of 21 surveys investigating research misbehaviour found that about 2% of researchers admit to falsification or fabrication. About 34% of participants admitted to Questionable Research Practices (QRP) (4). QRPs embody a large class of research misbehaviours. However, since 14 of the 21 studies included in the meta-analysis focused on biomedical researchers, it is unknown if these rates generalise to other disciplinary fields. Similarly, when pooling the results of 17 studies investigating plagiarism, 1.7% of participants admitted to having plagiarised (13). However, 10 of those studies used a biomedical sample. Hence, these results may not generalise across all sciences and the humanities. This also begs the question whether the research misbehaviours that participants were asked about are actually relevant to their own research, as some QRPs may be field or even discipline-specific.

We investigated whether the importance of particular research misbehaviours, as perceived by researchers in their respective disciplines, varies across disciplinary fields.

In our study, we distinguish four major disciplinary fields: biomedical sciences, natural sciences, social sciences and the humanities. Our study aims to assess what academic

researchers in Amsterdam consider to be the most important research misbehaviours in their respective disciplinary fields.

Methods

Design

We used a mixed methods sequential explanatory quantitative first design (14). This implies that our study had two phases: 1) a quantitative phase in which we collected survey data, and 2) a qualitative phase in which we conducted focus group interviews to deepen our understanding of the survey responses. See figure 1.

Figure 1. Overview of study design and analysis.

Ethical statement

The Scientific and Ethics Review board of the Faculty of Behavioural and Movement Sciences, Vrije Universiteit, Amsterdam reviewed and approved our study.

Participants

Participants consisted of academic researchers with at least a 0.2 FTE research appointment at the Vrije Universiteit Amsterdam, University of Amsterdam or the Amsterdam University Medical Centers and included PhD candidates.

100 *Quantitative data collection procedure*

101 We contacted the deans and rectors from the participating institutions with a request
102 to contact their academic researchers. The institutions shared the contact details of their
103 researchers on the basis of a formal data sharing agreement. To explain the study's aim, we
104 sent all academic researchers in Amsterdam ($n = 7548$) an information letter. This letter also
105 included a hyperlink to the privacy policy and the study protocol on our project website (see
106 appendix 1 and 2 and www.amsterdamresearchclimate.nl). One week later we sent an
107 invitational email to all researchers. Participants had to give informed consent and confirm
108 that they were actually involved in research for at least one day per week on average
109 (inclusion check) at the beginning of the survey. We used Qualtrics (Qualtrics, Provo, UT,
110 USA) to build the survey.

111 Participants were randomly presented 20 major and minor research misbehaviours
112 from a list of 60 (see Bouter et al (10) for an overview and appendix 3). We thus used a
113 planned missing data design as a means to shorten the overall questionnaire length and to
114 reduce the risk that participants get tired and quit the survey (15).

115 For each research misbehaviour, participants first had to indicate the perceived
116 frequency ("How often have you observed the behaviour stated above in the last three
117 years?"). Answer options were 1 ("Never"), 2 ("Once or twice") and 3 ("Three times or
118 more"). This question had to be answered in reference to their main disciplinary field.
119 Second, participants were asked to assess the impact of the research misbehaviour on the
120 validity of the findings ("If you were to observe this behaviour, how large would its impact
121 be on the validity of the findings of the study at issue?") on a scale from 1 ("Negligible") to 5
122 ("Enormous"). Finally, the survey collected basic demographic data: participants' academic
123 rank (PhD student, postdoc, assistant professor, associate professor or full professor),

disciplinary field (biomedicine, natural sciences, social sciences and humanities) and gender (male or female). For analyses purposes, we clustered academic rank into three groups: PhD students, postdocs and assistant professors and finally associate and full professors. The survey contained two other instruments; the results of those have been presented elsewhere (16,17).

Quantitative data analysis

We preregistered our intended analyses on the Open Science Framework, see <https://osf.io/x6t2q/register/565fb3678c5e4a66b5582f67>. Here we explain the main analyses briefly. First, we calculated the five most frequent and five most impactful research misbehaviours per academic rank and disciplinary field. Second, to assess which research misbehaviours were most ‘important’, we computed aggregated impact scores with the mean product of frequency and impact. Importance is thus defined as the product of frequency and impact. It ranged from 0 (negligible impact/never observed this) to 15 (enormous impact/observed this more than three times). We present these stratified top 5 rankings of important research misbehaviours below.

Qualitative data collection

To better understand the relevance of the aggregated impact top 5 research misbehaviours established in the survey, we organized focus groups with researchers from three academic ranks and four disciplinary fields. These focus groups took place at the Vrije Universiteit, therefore we only invited researchers from the Vrije Universiteit and the Amsterdam UMC (location VUmc) as they are most conveniently located.

We recruited researchers in three ways. First, by writing to heads of department and asking them to provide e-mail addresses of potentially interested researchers. Second, via our

own network of colleagues that work on unrelated topics. Third, we randomly selected researchers from the different academic ranks and disciplinary fields and invited them via e-mail where we explained the purpose of the focus group and asked them to participate. When an invitee abstained from participation (abstaining from participation was mostly due to conflicting schedules, lack of time or other reasons), we invited a new researcher, and so on until we reached a minimum of 4 confirmations per focus group.

In total, between March 2018 and May 2018, we conducted 12 focus group interviews with 61 researchers. To encourage participants to speak freely, the groups were homogenous for academic rank and disciplinary field, see table 1.

Table 1. Overview of academic researchers from Vrije Universiteit Amsterdam and Amsterdam UMC location VUmc per focus group.

Academic rank	Disciplinary field*			
	<i>Biomedicine</i>	<i>Natural sciences</i>	<i>Social sciences</i>	<i>Humanities</i>
<i>PhD students</i>	5 (5)	4 (0) ^E	4 (3) ^E	6 (5) ^E
<i>Postdocs and assistant professors</i>	5 (4)	3 (0)	7 (3) ^E	5 (5) ^E
<i>Associate and full professors</i>	4 (0)	4 (0)	4 (1) ^E	7 (3) ^E
Total	14 (9)	11 (0)	13 (7)	18 (13)

* In brackets is the number of female researchers.

^E denotes focus groups that were conducted in English.

A facilitator (TH or JT) led the focus groups, accompanied by an observer who made notes and ensured audiotape recording. We constructed a topic guide to direct the focus group

interviews (see appendix 4) where we presented participants with the aggregated impact top 5 of research misbehaviours that we had found in the survey among researchers from their disciplinary field. We then asked participants to add new research misbehaviours that were, in their opinion, at least as important in their disciplinary field. As a restriction, we asked all researchers to focus on things they had actually experienced or observed, instead of something they had only heard of or read about.

We used Participative Ranking Methodology (18) where participants had to obtain consensus on which research misbehaviours they deemed most important. They did so by ranking the survey research misbehaviours and their own newly formulated research misbehaviours on a scale from major to minor. We took photos of the final rankings of research misbehaviours of each group. In total, this took between 20 and 35 minutes. The remaining results of the focus groups will be part of another report.

Qualitative data analysis

We read the transcripts and started open coding using Atlas TI© Version 8.3.0. If the transcripts were in Dutch, we assigned English codes for consistency and translated quotations. We used inductive content analysis to analyse the transcripts as it is a good method for systematically describing and understanding complex phenomena (19) and it helps to reduce rich data to meaningful concepts that capture the phenomenon of interest (20). We thus grouped codes to identify emerging themes. Team members (JT, TH, GW and RP) independently identified themes and these were discussed to achieve consensus and thereby increase reliability. These themes were organised in a code tree and applied to other transcripts (see appendix 5). Finally, we identified appropriate quotes to illustrate each theme.

Results

Ninety-two e-mail addresses were no longer in use and 146 researchers filled in the non-response questionnaire. Hence 7310 potential respondents were left of which 1080 researchers completed the 60 items. Survey completion rate was 15% (see figure 2). First, we present the quantitative top 5 of important research misbehaviours per disciplinary field. Second, we provide the relevant themes from the focus groups that shed more light on what these and other important research misbehaviours mean and illustrate these with quotes.

Figure 2. Overview of survey response rate.

Disciplinary fields

An overview of the top 5 most frequent and most impactful research misbehaviours per disciplinary field can be found in appendices 6 and 7. The top 5 important research

misbehaviours stratified per rank can be found in appendix 8. Finally, a stratified ranking of all 60 items can be found in appendix 9.

Briefly, the misbehaviour ‘fabrication of data’ qualified as most impactful for the validity of the study in all disciplinary fields. Most frequent research misbehaviours differed somewhat. Biomedical researchers perceived listing an author that doesn’t qualify for authorship to be most frequent. According to natural sciences researchers and social sciences researchers, insufficient supervision was most frequent. Researchers in the humanities perceived selective citation to be most frequent. Humanities researchers rated the presentation of grossly misleading information in a grant application as having most impact.

In this paper, we focus on the top 5 of most important (importance was defined as the product of frequency and impact scores) research misbehaviours per disciplinary field, see table 2.

Table 2. Top 5 important research misbehaviours by disciplinary field, based on the survey.

Top 5 important research misbehaviours per disciplinary field with <i>M (SD)</i>								
	Biomedicine		Natural sciences		Social sciences		Humanities	
#1	Insufficiently supervise or mentor junior co-workers	7.02 (3.63)	Insufficiently supervise or mentor junior co-workers	7.72 (4.13)	Insufficiently supervise or mentor junior co-workers	6.95 (3.78)	Insufficiently supervise or mentor junior co-workers	6.76 (3.84)
#2	Choose a clearly inadequate research design or using evidently unsuitable measurement instruments	6.04 (3.16)	Not report clearly relevant details of study methods	6.95 (3.43)	Not publish a valid ‘negative’ study	6.54 (3.98)	Use published ideas or phrases of others without referencing	6.69 (3.69)
#3	Let own convictions influence the conclusions substantially	5.99 (3.17)	Insufficiently report study flaws and limitations	6.64 (3.41)	Let own convictions influence the conclusions substantially	5.86 (2.95)	Selectively cite to enhance own findings or convictions	6.17 (3.25)
#4	Give insufficient attention to the equipment,	5.64 (3.32)	Let own convictions influence the	6.38 (3.27)	Choose a clearly inadequate research	5.77 (3.38)	Choose a clearly inadequate research	6.11 (3.37)

	skills or expertise which are essential to perform the study		conclusions substantially		design or using evidently unsuitable measurement instruments		design or using evidently unsuitable measurement instruments	
#5	Keep inadequate notes of the research process	5.62 (2.96)	Give insufficient attention to the equipment, skills or expertise which are essential to perform the study	6.26 (3.48)	Give insufficient attention to the equipment, skills or expertise which are essential to perform the study	5.71 (3.3)	Unfairly review papers, grant applications or colleagues applying for promotion	6.03 (4.15)

We presented this top 5 to the focus groups participants and asked them to reflect on these items and add any relevant research misbehaviour from their disciplinary field. They then had to rank the survey research misbehaviours and their personally perceived research misbehaviours from major to minor importance. These rankings can be found in appendix 10.

From our qualitative analysis, the majority of research misbehaviours fell into one of three broad categories: issues around peer review, sloppy conduct of research and insufficient supervision. To better understand what sort of research misbehaviours researchers from a particular disciplinary field were confronted with and how these impacted their research, we zoomed in on themes that were more specific for a disciplinary field or that received more attention in their discussions. We present these themes per disciplinary field and where possible, we identified quotes as illustrations, see table 3 below.

Biomedicine: delaying reviewers, sloppy reporting and insufficient supervision

Biomedical researchers were vexed about inflexible reviewers that either delayed the publications of their findings or that were unresponsive to valid counterarguments in rebuttal letters when the manuscript challenged the mainstream view in the field. This made it particularly hard to publish negative research findings, whereas focus group participants agreed that this was pivotal for knowledge to progress.

237 Another hindrance for knowledge to progress was that authors drew (wrong)
238 conclusions based on little solid argumentation or seemed to interpret the data as it suited
239 them. This was especially pertinent when only the most positive findings were reported, that
240 then led to replication problems because the positive result was likely obtained by chance.

241 Insufficient supervision was a concern that participants recognised but they also
242 indicated that a PhD student is expected to ask for help when in need. In addition, the
243 supervisor can make the PhD candidate aware of existing time pressures, but this should be
244 realistic and not indicate that the PhD student is not allowed to take holidays. Finally, it was
245 generally agreed upon that little supervision is not a sufficient condition for irresponsible
246 research, yet it could increase the chances of a PhD student conduct irresponsible research.

247

248 *Natural sciences: Review misconduct and no team spirit*

249 Natural scientists brought up the topic of review misconduct. The misconduct takes
250 the form of the editor or reviewer stealing ideas when put forth for publication or in a grant
251 proposal. Reviewers or editors could either postpone publication and quickly publish the idea
252 themselves or they could reject the manuscript and publish its ideas elsewhere. A similar
253 story holds for grant proposals.

254 Natural science researchers also noted that insufficient communication and
255 supervision may damage team relations and some researchers may fail to put their success
256 into context, claiming that the success is only theirs.

257

258 *Social sciences: sloppy reviewing, sloppy design and statistics and insufficient supervision*

259 Social science researchers often encountered reviewers that demanded to be cited,
260 which is obviously not the purpose of the review. Furthermore, they encountered reviewers
261 that were not up to scratch with developments in the field. Lastly, some had experience with

reviewers that failed to declare a conflict of interest as they had a previous relationship with the authors, revealing nepotism in publication review.

Another concern was the sloppy methods where researchers referred to conducting an underpowered study or failing to report a non-replication. Related was the use of ‘harking’ (hypothesizing after results are known), where supervisors encouraged their PhD students to present an unexpected finding as being hypothesised from the start (21). Other examples involved collecting more data when results were almost significant or just pressing PhD students into ‘finding’ an effect in the data, even when probably no actual effect was there.

Finally, concerns were voiced about insufficient supervision of PhD students. More senior researchers noted that PhD students were being held for their academic projects at a very early stage in their career, when a PhD student is still learning what academic research involves. Sometimes supervisors took advantage of their PhD students, either by demanding co-authorships without a justification or by mentally agitating their PhD students.

Humanities: uncritical reviewing, mediocre research and scarce supervision

One concern regarded uncritical reviewing. That could mean a reviewer reviewing without specific comments, or reviewers that just accept a paper because of the authority of the author. This could be due to the fact that peer reviewing is not valued high enough by the scientific community. Another form of uncritical reviewing regarded failure to filter out fake papers that were clearly a hoax. Participants connected this to the fact that some fields lack clear publication criteria that a reviewer can use to judge a manuscript’s potential.

A second worry regarded mediocre research, which could mean research that is not value free, opaque or hastily written up, repetitive and inflating small findings. A related research misbehaviour was the stealing of original ideas from colleagues but also stealing ideas from PhD or master students and publishing it without (even) acknowledging them.

Finally, humanities researchers noted scarce supervision could lead to fraud. ‘Scarce’ could be in terms of quantity; there are very few postdocs and hence there is no day-to-day supervision of PhD students. ‘Scarce’ could also regard the quality, such as when supervisors don’t take their responsibility or supervisors who are actually not an expert on the topic of the PhD student are assigned to be supervisor.

Table 3. Quotations per disciplinary field to illustrate the content of the research misbehaviour themes.

<i>Biomedicine</i>	
Theme	Quote
Sloppy reporting	“Take things that are reported as a decrease of 80% in tumour rate. Well, when you attempt to repeat the experiment you get a 60% decrease so obviously their 80% was the most positive result from all the times they tried...” – <i>PhD candidate</i>
Insufficient supervision	“If you have a PhD candidate and you completely throw her in at the deep end, surely you increase the chance of irresponsible research” – <i>Full professor</i>
Inflexible reviewers	“So everything that is novel or different, it requires an lot of effort to get that accepted in the, in the journals, due to most likely also rigid reviewers” – <i>Assistant professor</i>
<i>Natural sciences</i>	
Theme	Quote
Review misconduct	“I had it once with a journal editor who was being really difficult about a publication of mine. And then he managed to get his own publication [with the same idea] in before mine” – <i>Full professor</i>
Team spirit missing	“Research is no one man show, you have to teach them [PhD candidates] to also let go, it is not just theirs. The same holds for what I do, it is not just mine, it is a team effort...” – <i>Assistant professor</i>
<i>Social sciences</i>	
Theme	Quote
Sloppy reviewing	“You’re on a grant review panel and you’re judging someone whom you have a personal or professional relationship with. You’re an editor of a journal and you don’t recuse yourself for a conflict of interest with the author of a paper” – <i>Associate professor</i>
Sloppy methods and statistics	“What is so horrible about these strategies is, post-hoc storytelling, salami slicing, is how you win the game, this is how you become a professor, this what you should do. Some professors even tell you, like: this is what you should do” – <i>Postdoctoral researcher</i>
Insufficient supervision	“Supervisors exploiting their PhD students. I think that can be sort of extended into any type of harassment; sexual, personal, mental harassment, whatever it is. Also about any type of power relationship that there is and... that he demands co-authorships, that supervisors say... I want to be on this paper, I am on this paper, not as a question but, you know, as a statement...” – <i>postdoctoral researcher</i>
<i>Humanities</i>	
Theme	Quote
Uncritical reviewing	“What you see is that, there is no review culture, in which the standards of what constitutes good and bad publications are adequately present, to filter out actual hoaxes” – <i>Full professor</i>
Lack of supervision	“I have a PhD student who got sent to me from abroad... I said well, when did you last speak to your supervisor? And he said no, no, no, because you can answer my questions better, the last couple of months I didn’t, because I was saving it up for you... While the actual supervision who will get... the credits is actually not an expert.” – <i>Assistant professor</i>

295
296 * Brackets added by the authors.

297

298

299 **Discussion**

300 This study was based on a survey followed by focus groups to get insight in what
301 academic researchers in Amsterdam from different disciplinary fields considered to be the
302 most important research misbehaviours. Participants ranked insufficient supervision as most
303 important regardless of disciplinary field. The ranking of other misbehaviours differed
304 somewhat between disciplines: natural scientists assigned importance to failure to report
305 clearly relevant details of study methods, social scientists ranked not publishing negative
306 results highly and humanities researchers considered using ideas from others without
307 reference as important.

308 The focus group participants recognised the survey research misbehaviours as
309 important, but also brought a new understanding of the research misbehaviours to light that
310 they considered appropriate in their disciplinary field. Again, insufficient supervision as part
311 of poor mentorship was recognised as important by focus group participants in each
312 disciplinary field, although the item was not always classified as among the most important
313 research misbehaviours. Sometimes it was phrased differently: “Getting too little time for
314 research from mentor”. In addition, some misbehaviours seemed examples of what
315 insufficient supervision could involve, such as “Pressuring junior postdocs and PhDs into
316 ‘finding’ significant effects”. Nevertheless, these findings support the idea that insufficient
317 supervision is an important concern to academic researchers of all ranks and disciplinary
318 fields.

319 How focus group participants understood the research misbehaviours helps us
320 understand what these research misbehaviours actually mean to researchers. When comparing

what focus group participants from the biomedical and social sciences brought up, it may seem as if the research misbehaviours they discussed were already mentioned on the list of 60 used in the survey. However, different interpretations are possible here. First, it may be that the researchers filling in the survey *understood* or *interpreted* the presented research misbehaviours differently. For example: the survey item “Present grossly misleading information in a grant application” could be understood to include items that were brought upon in the focus group such as overpromising to funders, but it could also be that words like *grossly misleading* led survey participants to disregard overpromising.

Second, it could be that although similar at first sight, the research misbehaviours that the focus group participants discussed are actually different from the ones on the list of 60. For example, one minor misbehaviour from the survey was “Not report clearly relevant details of study methods”. An item mentioned in the focus group that may seem similar was “Not being transparent in reporting”. However, transparent reporting arguably involves the entire paper, not just the study methods section, hence this could genuinely be a different thing. Future research would have to investigate which research misbehaviours are two descriptions of the same phenomena, or two different phenomena.

Nevertheless, especially researchers from the natural sciences and humanities mentioned research misbehaviours that seemed missing or at least substantially different from the list of 60, such as referees or editors that abuse their power to steal original ideas. Future research will have to establish the frequencies, relevance and importance of the new items for these disciplinary fields. This would require translation the qualitative data into items and a representative sample from all disciplinary fields and is beyond the scope of the current study. To facilitate such an attempt, we provide an updated list of research misbehaviours (appendix 11) in which items are reformulated, included as explanatory examples or added as a new research misbehaviour.

346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370

Strengths

Our study may be the first that investigates research misbehaviours and includes researchers from different disciplinary fields and all academic ranks. It is noteworthy that the different methods we used (quantitative survey and qualitative focus groups) led to similar results as both survey and focus group participants recognised the importance of responsible supervision.

Additionally, our quantitative results largely confirm the previous findings by Bouter and colleagues (10). Their population consisted of visitors of the World Conference of Research Integrity, but apparently both groups identify insufficient supervision and sloppiness (haste) as important problems in contemporary academia.

Limitations

There are some study limitations to bear in mind. We had considerable non-response. However, response bias is not a necessary consequence of a low response rate as long as the respondents are representative for the population (22). Response bias occurs when responders differ with respect to the phenomena studied from non-responders. In our case, response bias may be bidirectional: Perhaps responders were more concerned about research misbehaviour than non-responders, this could lead to an overestimation. It could also be that non-responders witnessed research misbehaviours but did not feel protected enough to report these in our study, which would lead to an underestimation.

A further limitation is that our stratified top 5 of most frequent, impactful and important research misbehaviours do not represent absolute rankings. A brief inspection of the standard deviation leads us to conclude that what is listed as number 1 is not significantly more important than the misbehaviour on number 2. Nevertheless, the mean scores per item

371 decrease after the top 5 (see appendix 9), so the fact that an item is in the top 5 is a robust
372 indicator for its importance, albeit not for its exact rank number.

373 Another limitation is that disciplinary fields are heterogeneous and overlapping. A
374 more precise phrasing such as keeping in mind one's research department may have provided
375 more valid frequency estimates. Yet, in that case multiple respondents may base their
376 estimates on the same cases, leading to duplication.

377 Finally, the participatory ranking exercise may have been difficult for participants.
378 Especially when trying to determine whether a research misbehaviour is major or minor, it
379 was sometimes discussed that a research behaviour can have of major impact on an
380 individual researcher but relatively minor impact on academic research as a whole.
381 Notwithstanding, the focus group participants eventually agreed on how to rank the research
382 misbehaviours, which often involved a combination of impact on the individual and impact
383 on research at large.

384

385 *Implications*

386 Since our findings reproduced previous ones in that insufficient mentoring was
387 recognised as an important research misbehaviour regardless of disciplinary field, we think
388 that exploring interventions that foster responsible mentoring are worthwhile. Possibly
389 interventions that optimise mentoring skills should go beyond essential soft skills and include
390 reflection on how new researchers can best navigate in a research environment where
391 research misbehaviour may occur. Once tested for effectiveness, these supervision trainings
392 could pave the way for a more responsible research climate.

393

394 **Conclusion**

Academic researchers across disciplinary fields consider insufficient supervision and sloppy science the most important research misbehaviours. A few discipline-specific QRPs were found. In particular, in the natural sciences and humanities the abuse of referee or editor power by stealing ideas from proposals or manuscripts they have to evaluate are important. Our findings underscore broad communalities and a few differences between disciplinary fields. Yet, future research should investigate what respondents understand when scoring the research misbehaviours. Development of programmes to incentivize and optimise supervision of junior co-workers should be prioritized in academia.

List of abbreviations

QRP: Questionable Research Practices

FTE: Full Time Employment

OSF: Open Science Framework

Declarations

Ethics approval and consent to participate

Our study was ethically approved by the Scientific and Ethical Committee of the Faculty of Behavioural and Movement Sciences, Approval Number: VCWE-2017-017R1.

Consent for publication

Not applicable.

Availability of data and material

416 The datasets generated and/or analysed during the current study are not publicly available due
417 participants' privacy but are available in pseudo-anonymised version from the corresponding
418 author on reasonable request under a data sharing agreement.

419

420 *Competing interests*

421 The authors declare that they have no competing interests

422 *Funding*

423 LB, JT and TH were partly supported by the Templeton World Charity Foundation
424 (<https://www.templetonworldcharity.org/>) under the grant #TWCF0163/AB106. Additional
425 support was provided by the University of Amsterdam, the Vrije Universiteit and the
426 Amsterdam University Medical Centers. The funders had no role in the design of the study
427 and collection, analysis, and interpretation of data and in writing the manuscript

428 *Authors' contributions*

429 LB, GW, JT, GtR, RP and TH designed the study. TH and JT collected and analysed the data.
430 TH wrote the draft. LB, GW, JT, GtR, RP commented several times and the author team met
431 to discuss the core issues. All authors read and approved the final manuscript.

432 *Acknowledgements*

433 We would like to thank all members of the steering committee of the Academic Research
434 Climate Amsterdam project (professor F. Oort, professor Y. Smulders, professor H. de Haes
435 and professor R. van Woudenberg) for their critical comments during presentations on the
436 study design or draft communications regarding the study or interpretation of the data.

References

1. Veldkamp CLS, Hartgerink CHJ, van Assen MALM, Wicherts JM. Who Believes in the Storybook Image of the Scientist? *Account Res.* 2017;24(3):127–51.
2. Mahoney MJ. Review Paper: Psychology of the Scientist: An Evaluative Review. *Soc Stud Sci.* 1979;9(3):349–75.
3. Mahoney MJ. *Scientist as subject: The psychological imperative.* Cambridge, MA SE: Ballinger Publishing Company; 1976.
4. Fanelli D. How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS One.* 2009;4(5):e5738.
5. Martinson BC, Anderson MS, de Vries R. Scientists behaving badly. *Nature.* 2005;435(7043):737–8.
6. Kornfeld DS. Perspective: Research misconduct: The search for a remedy. *Acad Med.* 2012;87(7):877–82.
7. Godecharle S, Nemery B, Dierickx K. Guidance on research integrity: No union in Europe. *Lancet.* 2013;381(9872):1097–8.
8. Bedeian A, Taylor S, Miller A. Management science on the credibility bubble: Cardinal sins and various misdemeanors. *Acad Manag Learn Educ.* 2010;9(4):715–25.
9. Bouter LM. Commentary: Perverse incentives or rotten apples? *Account Res.* 2015;22(3):148–61.
10. Bouter LM, Tjldink J, Axelsen N, Martinson BC, ter Riet G. Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity. *Res Integr Peer Rev.* 2016;1(17):1–8.
11. Steneck N. Fostering integrity in research: Definition, current knowlege, and future directions. *Sci Eng Ethics.* 2006;12(1):53–74.
12. De Vries, Raymond; Anderson, Melissa; Martinson B. Normal Misbehavior: Scientists

- Talk About the Ethics of Research. *J Empir Res Hum Res Ethics*. 2006;1(1):43–50.
13. Pupovac V, Fanelli D. Scientists Admitting to Plagiarism: A Meta-analysis of Surveys. *Sci Eng Ethics*. 2015;21(5):1331–52.
14. Creswell JW. *Research design: Qualitative, quantitative, and mixed method approaches*. 4th ed. Los Angeles: Sage; 2014.
15. Graham JW, Taylor BJ, Olchowski AE, Cumsille PE. Planned Missing Data Designs in Psychological Research. *Psychol Methods*. 2006;11(4):323–43.
16. Haven TL, Tijdink JK, Martinson BC, Bouter LM. Perceptions of research integrity climate differ between academic ranks and disciplinary fields: Results from a survey among academic researchers in Amsterdam. *PLoS One*. 2019;14(1):e0210599.
17. Haven TL, Bouter LM, Smulders YM, Tijdink. Perceived publication pressure in Amsterdam –survey of all disciplinary fields and academic ranks. *Prepr ahead Publ*. 2018;
18. Ager A, Stark S, Potts A. *Participative ranking methodology: A brief guide*. Mailman School of Public Health. 2010.
19. Krippendorff K. *Content analysis: An introduction to its methodology*. Beverly Hills SE: Sage Publications; 1980.
20. Elo S, Kyngäs H. The qualitative content analysis process. *J Adv Nurs*. 2008;62(1):107–15.
21. Kerr NL. HARKing: Hypothesizing After the Results are Known. *Personal Soc Psychol Rev*. 1998;2(3):196–217.
22. Cook C, Heath F, Thompson R. A meta-analysis of response rates in web-or internet-based surveys. *Educ Psychol Meas*. 2000;60(6):821–36.

Appendices

487 Appendix 1. Privacy policy

488 Appendix 2. Research protocol

489 Appendix 3. Full list of 60 items

490 Appendix 4. Focus group topic guide

491 Appendix 5. Code tree

492 Focus group themes are displayed colour codes where pink is social sciences, purple is

493 humanities, blue is natural sciences and green is biomedical sciences. Higher order

494 themes are dark purple and connect to the overall theme, which is light green. Lines

495 signify relations (i.e. is one example of/is part of).

496 Appendix 6. Top 5 most frequent research misbehaviours, per disciplinary field and academic

497 rank.

498 Appendix 7. Top 5 most impactful research misbehaviours, per disciplinary field and

499 academic rank.

500 Appendix 8. Top 5 most important research misbehaviours per academic rank.

501 Appendix 9. 60 items ranked for importance per disciplinary field and academic rank.

502 Research misbehaviours that appeared in the top 5 of that particular subgroup are

503 printed in **bold**. M = mean score per subgroup, SD = standard deviation, N = number

504 of responses. Note that we used a missingness-by-design, so the actual number of

505 respondents from that particular subgroup should be multiplied by three. Importance

506 was computed as the product score of frequency (1-3) and impact (1-5) and thus

507 ranged from 1 to 15. The higher the mean score, the more important this research

508 misbehaviour was perceived according to researchers from a particular subgroup.

509 Appendix 10. Qualitative rankings of important research misbehaviours

510 Appendix 11. List of 60 misbehaviours with alternative formulations and newly added items

511 from the focus group interviews printed in *italics*. Newly added and preferred

512 formulations of existing items are printed ***italic-bold*** and the former formulation is
513 ~~crossed out.~~

Appendix 1. Privacy Policy Academic Research Climate Amsterdam

Privacy Policy - Project Academic Research Climate in Amsterdam (ARCA)

Introduction

This privacy policy informs participants how we comply with the privacy regulations of the participating institutions. We have received your email address from the Executive board of one of the four academic institutions in Amsterdam (VU, VU University Medical Center, University of Amsterdam and Academic Medical Center). With the boards of these institutions, we have made a data transfer agreement to obtain the email addresses of all active scientists. In this agreement we are committed to the privacy laws of the four institutions. The agreement stipulates that we will code all data after sampling and anonymize all obtained data before analysis. The institutions will have no access to the research data. They will only be informed with aggregated reports on a faculty level as specified below.

Aim of the study

The goal of the ARCA project is to map the research climate within the four academic institutions in Amsterdam.

Ethical considerations

The ARCA processes, manages and secures personal information with the utmost care. We work according to the Dutch Code of Conduct for Scientific Practice.

Therewith we comply with all the requirements of the Dutch Personal Data Protection Act. The study protocols were reviewed by the Medical Ethics Committee (METc) of the VU University Medical Center. The METc reviewed the protocol and stated that the study does not fall under the scope of the Medical Research Involving Human Subjects Act (WMO). The study protocol was reviewed and has been approved by the Ethics committee of the VU, faculty of Behavioral and Movement Sciences.

Scope of this Privacy Policy

This privacy policy applies to all personal information we collect as part of the ARCA project. After data collection, we will de-couple email address from responses after which email addresses will be deleted from the database before analysis and aggregation of data. Only data on age categories (by ten years), academic rank, gender and disciplinary field (fields are categorized into four types: biomedicine, social sciences, physics/engineering and humanities) will be used. These procedures

described here will effectively anonymize the data to assure that it will be practically impossible to trace back personal responses to specific questions.

Provision of anonymized data to third parties

The de-identified data (raw data after removal of email address) will not be provided to third parties unless:

- This is necessary for the purpose for which the data were obtained (i.e. submission of the de-identified dataset as part of publication policies).
- When our anonymized data is needed for research purposes in the future (i.e. a collaboration with another institution), your identity will remain protected and data will only be available after a data transfer agreement is signed by the requesting party and the researchers.
- When such a request for the anonymized data is issued, data will only be shared if the main research team reaches consensus that the requester's use of the data for future research is strictly necessary.
- If any other institution or journal wishes to us to make our data available, this will only be done via a data-transfer agreement.

Data protection

We use security and safety procedures against theft, loss and abuse by third parties and ensure that only authorized persons have access to the database. Authorized persons are the members of the research team. The research team comprises the principal investigators (J. Tjldink, MD PhD, Prof. L.M. Bouter PhD, and T.L. Haven, MSc) and the statistician of the VU / VUmc.

Data storage and location

We will store your data for a maximum of fifteen years. Data for scientific research is stored according to the legal timeframe of data storage. The data is collected through Qualtrics, which means that the data is stored on an European server.

Dissemination of results for research

We will analyze and publish the results of the survey for research purposes. Therefore we will not include data on faculty, institution or university but will only stratify the data for disciplinary fields and academic ranks.

Dissemination of results for institutions

The reports for the four institutions will be aggregated at faculty level. Because universities are split into different faculties, they will receive reports on a faculty level. Medical centers are split into research institutes and will receive one report per institute. Summaries are only written if more than 25 active scientists filled in the survey and are working at the faculty. This will make it practically impossible to identify individual participants.

Questions

If you have any questions, please send an email to Tamarinde Haven, department of philosophy via info@amsterdamresearchclimate.nl. Your query will be dealt with as quickly as possible and will remain confidential.

Withdrawing from the study

If you would like to withdraw from the study, please send an email to Tamarinde Haven, department of philosophy via info@amsterdamresearchclimate.nl to complete withdrawal.

Academic Research Climate in Amsterdam study protocol

Table of contents

Summary

1. **Project history**
2. **Research Team**
3. **Background**
4. **Research questions**
 1. **Quantitative part**
 2. **Qualitative part**
 3. **Pilot interventions**
1. **Methods used**
 1. **Ethical considerations**
 2. **Survey**
 - 2.1. **Who?**
 - 2.2. **How?**
 - 2.3. **When**
 - 2.4. **Preparations**
 3. **Focus group interviews**
 - 3.1. **Who?**
 - 3.2. **How?**
 - 3.3. **When**
 - 3.4. **Preparations**
 4. **Pilot interventions**
 - 4.1. **Who?**
 - 4.2. **How?**
 - 6.4.2.1A Moral case deliberations
 - 6.4.2.1B Education program: mentoring the mentor
 - 4.3. **When**
 - 4.4. **Preparations**

6. **Relations of this protocol with the TWCF project**
7. **Envisioned output**
8. **Timeline**
9. **Budget**

Summary

Ensuring research integrity should be one of the key responsibilities for universities and university medical centers. Traditionally, institutions have focused on codes of conducts, procedures for handling allegations and courses on responsible conduct of research. Recently the importance of fostering a climate of research integrity receives more attention. The aim of our project is to explore and clarify the most salient aspects of the research climate in the four academic institutions in Amsterdam and to identify promising ways to promote responsible research practices. Special attention will be given to differences between academic ranks and disciplinary fields. We will assess the salient aspects of research climate by a web-based survey among all active researchers ($n > 5000$) using the Survey of Organizational Research Climate. The level of perceived publication pressure will be determined with the Publication Pressure Questionnaire. We will also explore the perceived frequency of occurrence of 60 research misbehavior items, what their impact is on validity of study findings and explore to what extent these misbehaviors impact the trust in science. Departing from the survey results we will determine with focus group interviews what the perceived barriers for responsible conduct of research are. Also the thoughts and perceptions of scientists on the preventability of research misbehaviors will be explored as are the solutions and interventions believed to be most effective. Subsequently we will design and try out two pilot interventions that are based on the knowledge gathered in the web-based survey and the focus group interviews: 1) regular moral case deliberation sessions in research groups, and 2) a training program for novice mentors of PhD students.

1. Project history and formation

The project Academic Research Climate in Amsterdam is a merger of two different research proposals that complement each other. The first part of the larger project called The Epistemic Responsibilities of the University, which has been granted by the Templeton World Charity Foundation (TWCF - see **Annex 1** for full description of this project). In the part of the TWCF project that is contained in this study protocol, we collect quantitative and qualitative data to investigate what aspects of the research climate promote or hinder research integrity, specifically focused on the Vrije Universiteit Amsterdam. We will develop two pilot interventions that aim to foster responsible conduct of research.

The second proposal (See **Annex 2**) that is merged in this study protocol is an extension of the TWCF project (**Annex 1**) with co-financing from the institutions collaborating in the Amsterdam Academic Alliance (AAA) and is called Research Culture in Amsterdam (**Annex 2**). This project has the same research questions as the relevant part I of the TWCF project – with the exception of the piloting of two interventions – and is in fact its extension to the three other academic institutions in Amsterdam (VU University Medical Center, Academic Medical Center and University of Amsterdam).

2. Research Team:

Core team

Lex Bouter, PhD, Professor of Methodology and Integrity, Department of Philosophy, Faculty of Humanities, Vrije Universiteit and Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam

Joeri Tjldink, MD, PhD, postdoctoral researcher and psychiatrist, Department of Philosophy, Faculty of Humanities, Vrije Universiteit Amsterdam

Tamarinde Haven, PhD student, Department of Philosophy, Vrije Universiteit Amsterdam

Steering Committee

Rene van Woudenberg, PhD, Professor of Epistemology and Metaphysics, Department of Philosophy, Faculty of Humanities, Vrije Universiteit Amsterdam

Yvo Smulders, MD, PhD, Professor of Internal Medicine, VU University Medical Center

Gerben ter Riet, MD, PhD, Associate Professor of Epidemiology, Amsterdam Medical Center

Hanneke de Haes, PhD, Professor emeritus of Medical Psychology, Amsterdam Medical Center

Frans Oort, PhD, Professor of Methods and Statistics, University of Amsterdam

Guy Widdershoven, PhD, Professor of philosophy and medical ethics, Vrije Universiteit, Amsterdam

Brian Martinson, PhD, Senior Research Investigator, HealthPartners Institute for Education and Research, Minneapolis, Minnesota, USA

3. Background

Optimizing the validity and reproducibility of the results of scientific investigations is one of the key responsibilities of universities and university medical centers. Traditionally, universities have served these purposes by conventional efforts, such as promoting methodological rigor and rewarding successful scientists. Lately, however, it has become clear that it is also necessary to specifically foster research integrity by the promotion of responsible conduct of research (RCR) and the prevention of research misconduct (RM) and questionable research practices (QRP). To realize these aims, various interventions are available: offering education, providing facilities, issuing rules, codes, and guidelines, or influencing the research climate directly. Unfortunately, very little sound empirical evidence is available as to what the most important issues are. Also little is known about the views of scientists regarding the way the research climate promotes or hinders research integrity. This project will help to collect this evidence. We will do this within and to the benefit of our own academic institutions, but with a view to also generate generalizable insights. The main research question is:

What do scientists of the four academic institutes in Amsterdam consider to be the salient aspects of the research climate of their institution that promote or hinder research integrity, and which do they believe to be the most important barriers to responsible conduct of research (RCR) and the most promising interventions to prevent research misbehavior (RM) and questionable research practices (QRP)?

We will collect empirical evidence in the four academic institutions that collaborate in the Amsterdam Academic Alliance (AAA): University of Amsterdam, Vrije Universiteit, VU University Medical Center, and Amsterdam Medical Center.

Special attention will be given to differences between disciplinary fields (biomedical sciences, natural sciences, social sciences, and humanities) and between academic ranks (PhD students, postdoctoral fellows, assistant professors, associate professors, and full professors). Focusing the data collection on the four Amsterdam academic institutions enables intense interaction with the corresponding academic communities and an opportunity to identify promising interventions. We will analyze and report our data overall and separately for the four main disciplinary fields and academic ranks homogenous for 3 academic ranks (PhD students, postdocs plus assistant professors, and associate plus full professors). We will not do so for institutions, faculties, divisions, research institutes and departments. We will analyze similarities and differences between our data and the available findings with the questionnaires in other countries and compare the ranking results with the first study on ranking research misbehaviors (Bouter, Tjldink, Axelsen, Martinson, & ter Riet, 2016).

4. Research questions

The project will consist of three parts in which complementary parts of the main research question will be answered.

5.1 Quantitative part: web-based survey

- a. What are the perceptions of academic researchers in Amsterdam regarding the organizational climate for responsible research practices both in one's general organizational setting and in one's specific department, stratified for academic rank and disciplinary field?
- b. What is the level of publication pressure in the four academic institutions in Amsterdam, stratified for academic rank and disciplinary field?
- c. What do scientists in the four academic institutions in Amsterdam consider to be the most detrimental research misbehaviours, stratified for academic rank and disciplinary field?

2. Qualitative part: focus group interviews

- d. What are the perceived barriers for responsible conduct of research in different disciplinary fields?
- e. What are the experiences and views of scientists on preventability of questionable research practices in different disciplinary fields?
- f. What type of solutions and interventions can be identified that scientists expect to be the most effective, stratified for disciplinary field?

5.3 Part 3: Pilot interventions

- g. How can moral case deliberations foster responsible conduct of research?
- h. How can an educational program for novice mentors of PhD Students foster RCR?

6. Methods

The methods of the project are differentiated into three parts introduced above. The first part (the survey) will take place in the first year of the project. During this first year we will also start the preparations for part two and three: the focus groups at the different academic institutions that will be held in the second year of the project, and the pilot interventions in the third year of the project. Furthermore, during the first months of the project most energy and time will be used for getting formal endorsement by the four universities, the email-addresses of active scientists, and ethical approval). The second year of the project will consist of performing, analyzing and reporting focus group interviews. In the third year we will organize the two pilot interventions. We will try to gain insight in their practical implementation and effect by collecting information at baseline and measure same variables after one year.

6.1 Ethical considerations

All participants receive an information letter by mail prior to the survey invitation. In the invitation, a link to start the actual survey is contained. The first survey item asks the respondent to actively consent by indicating that they have understood the purpose of the study and that they voluntarily take part. The study will be reviewed by two independent ethics committees from two different institutions to determine its approval. The protocol will be submitted to the Institutional Review Board (METC) of the VUmc with a view to confirm that the Medical Research Involving Human Subjects Act (WMO) does not apply and to judge whether the proposed study is compliant with the pertaining privacy legislation. Furthermore the protocol will be reviewed by the ethics committee for behavioral and movement sciences from the VU University. Confidentiality is maintained using restricted, secure access to the data, destruction of audiotapes after transcription, and anonymous analysis of transcripts.

6.2. Web-based survey

6.2.1. Who?

TH will conduct the web-based survey under close supervision of JT. There will be a weekly meeting with TH and JT to discuss progress and for supervision. Furthermore we will have a fortnightly meeting with LB to determine progress, methods, drawbacks and results. Furthermore there will be regular (six times per year) meetings with the full steering group to discuss methods, limitations, progress and results.

6.2.2. How?

In the first year of the project we will conduct the web-based survey to get insight in what the members of our academic community believe to be the most salient aspects of the research climate and the most important ways to improve it.

The core of the questionnaire consists of two validated instruments: the Survey of Organizational Research Climate (Martinson et al. 2013; Crain et al. 2013; Wells et al. 2014) and the Publication Pressure Questionnaire (Tijdkink et al. 2013; Tijdkink et al. 2014; Tijdkink, Verbeke & Smulders 2014). (We have plans to improve the PPQ by re-analyzing the data and adding additional items to the questionnaire to increase construct validity.)

Additionally we will ask our respondents to rank major and minor research misbehaviors on frequency, impact, and preventability. This instrument was piloted among and discussed with 34 experts at the fourth World Conference on Research Integrity in June 2015 and surveyed among every WCRI participant of the past four World Conferences on Research Integrity (Bouter et al., 2016). Taken together the web-based survey will provide insight in how our academic community perceives a) research climate, b) publication pressure, and 3) research misbehavior. We will present the results stratified for academic rank (PhD students, postdocs & assistant professors, associate professors and full professors) and disciplinary field (biomedical sciences, natural sciences, social sciences and humanities). We will also summarize the data for each of the university faculties and the different research institutes of the university medical centers, (see table 1) , together with the corresponding overall results as a benchmark. This will enable informed discussions with a view to raise awareness and to identify actions that can be taken in the faculty or division at issue. We will only write a report per faculty/division if at least 25 respondents from that faculty/division took part in the survey to protect the participants' identity. We will invite all members of the four academic institutions in Amsterdam that are active in research (>0.2 fte) to participate in the survey. This large number of invitees (approximately 5,000) enables the combination of identity protection and detailed subgroup analyses. We also can afford to use a 100% sample because the costs of a web-based survey hardly depend on the number of participants. All invitees will get an e-mail with a link to the web-based survey. The invitational e-mail (See **Annex 3**) will also contain a link to decline participation. The latter will open a screen with a few simple questions that will enable a non-response analysis. We will formulate a strict privacy policy and provide a link to it in our invitational e-mail (**Annex 3**). Qualtrics is chosen as the online distributor of the web-based questionnaire as they have the most efficient and state of the art response techniques available. It is a professional and reliable third party that can guarantee anonymity.

For the research team it will be impossible to link records to e-mail addresses of respondents. Please note that the draft questionnaire (**Annex 4**) contains all 60 items of minor and major research misbehaviors, while each respondent will only get a random sample of 20 of these in random order. We will end the survey with demographic information on gender, academic position, faculty/division, and disciplinary field. We will aim to have a mean maximum time to complete the survey of 12 minutes.

We will use advanced marketing techniques to determine whether the invitational email is received, opened and whether the participant followed the link to the online survey.

A detailed data-analysis plan will be ready before the start of data collection (**Annex 12**).

UvA	VU	AMC	VUmc
-----	----	-----	------

Economics and Business	Behavioral and Movement Sciences	Cardiovascular Diseases	EMGO Intitute for Health & Care Research
Faculty of Dentistry (ACTA)	Faculty of Dentistry (ACTA)	Epidemiology and public health	IcarVU, Institute for Cardiovascular Research
Faculty of Humanities	Faculty of Humanities	Gastrointestinal Diseases and metabolic disorders	MOVE Research Institute Amsterdam
Faculty of Law	Faculty of Law	Infection and Immunity	Amsterdam Neuroscience
Faculty of Science	Faculty of Sciences	Neurological and Psychiatric Disorders	VUmc Cancer Center Amsterdam
Faculty of Social and Behavioral Sciences*	Faculty of Social Sciences	Oncology	Division VI
	Faculty of Theology	Reproduction and development	
	Economics and Business Administration		

*Table 1. Faculties (UvA and VU) and research institutes (AMC and VUmc) of the four academic institutions in Amsterdam. For more information on faculties in UvA and VU, see [link](#) & [link](#), on research institutes in AMC and VUmc, see [link](#) & [link](#) *Faculty of Social and behavioral Sciences consists of 3 different research institutions with a considerable number of scientists that may lead to three different reports.*

6.2.3. When?

The first part will take place in the first year of the project. See project timeline (**Annex 9**) for more details on planning.

6.2.4. Preparation

List of necessities for the start of the survey:

1. Access to all email addresses of all active scientists within the four institutions.
2. Permission of the rector's offices of the Vrije Universiteit (Prof. V. Subramaniam) and University of Amsterdam (Prof. K. Maex) and dean's offices of the VUmc (Prof. C. Polman) and AMC (Prof. J.H. Romijn) and from the ethical committees of the VUmc and by the ethics committee for behavioral sciences (the Amsterdam Institute of Social Science Research) from the university of Amsterdam (UvA) that performs an ethical review of our protocol to start the research project
3. Prepared invitational emails
4. Web-based survey (using Qualtrics) containing the Survey of Organisational Research Climate, the Publication Pressure Questionnaire, the ranking of research misbehaviors and relevant characteristics of the participants.

6.3 Qualitative part: Focus group interviews

6.3.1. Who?

The focus group interviews will be performed by different members of the research team. TH and JT will lead the group discussions when he/she has followed the necessary training in qualitative research methods. This can be organized in the first year of the project. JT has experience in conducting and analyzing focus group interviews and will supervise TH. The results will be analyzed by 3 scientists with experience in focus group study analysis to make sure the results are reliable and consistent.

2. How?

Participants

In the second year of the project we will conduct focus groups interviews (Creswell 1998; Krueger & Casey 2000; Tjldink et al. 2016) with members of our academic community to explore their views on the research culture in general and on the most important forms of RM and QRP in particular, as well the values that underlie these views. We will start from what was learned through the web-based survey, in order to deepen our understanding of what the results of the web-based survey really mean and imply for actions to be taken. The interviews will aim specifically at identifying perceived barriers to RCR and differences between academic ranks.

We will allow for differences between disciplinary fields. For that purpose we will conduct separate focus group interviews for biomedical sciences, natural sciences, social sciences, and humanities. Per disciplinary field we will first conduct 3 groups homogenous for academic rank (PhD students, postdocs plus assistant professors, and associate plus full professors), then 2 groups with mixed composition with a focus on the differences in views between the academic ranks that we identified in the homogenous focus groups. Finally we will conduct 2 focus group interviews with mixed composition of disciplinary fields but with a focus on homogeneous academic rank to determine differences among disciplinary fields. We will avoid having close colleagues in the same focus group. Together there will be $(3 \times 4 + 2 \times 4 + 2 \times 3) = 26$ focus group interviews with 8 participants each. By means of intensive discussion among the participants in the focus groups, views, thoughts and opinions of researchers will be collected. Scientists are eligible to participate if they are able to speak Dutch or English, are scientifically active (they should spend at least 0.2 FTE on research) and are willing to give informed consent. Scientists are recruited with the help of the rectors' and deans' offices of the faculties (VU and UvA) and divisions (VUmc and AMC) of the participating four institutions, each of which provided the email addresses of all active scientists. We will use random selection software specially designed for this project to randomly select the participants across the different disciplines and academic ranks and sent an invitation by email explaining the purpose of the focus group interviews. This random selection is essential to prevent selection bias. If the invited participant declined participation, we randomly select a second participant of the same type from the same department, and so on, until we have 6–8 participants from different groups/disciplines/departments per focus group.

Data collection and procedure

The research team will formulate possible discussion themes about research culture beforehand that will be based on our previous survey and a pilot version of a focus group interview that we conduct with a mix of fellow scientists from the research departments of the steering committee (Philosophy, Epidemiology/Biostatistics, internal medicine). The focus group interviews will last approximately 1.5–2 hours until the point when no new or relevant information emerge (attainment of saturation). The focus groups will be chaired by an experienced group leader that will guide the discussion (JT, TH or other experienced focus group leader). Each session will be audiotaped and transcribed verbatim. After the transcription, the material will be thoroughly analyzed independently by 2-3 experienced investigators from the research team (TH, JT and TBD) with expertise in qualitative research to get in depth information of the participants' views and opinions (Meadows & Morse 2001; Mays & Pope 2000). The analyses will then be compared and discussed with the other assessors (JT, TH and 3

independent researchers) to formulate qualitative conclusions on the identified themes and to choose exemplary quotes. The results from the focus group interviews homogenous for academic ranks will inform the interview schemes for the focus group interviews with mixed composition.

A semi structured protocol (**Annex 6**), partially based on the pilot version of the focus group will be used as a structured guidance for conducting the focus group interviews. This includes information on general aspects of focus groups, an introduction to the subject, and an initial exploration of the participants' first thoughts on research integrity to increase engagement with the discussion. After this, participants will be invited to present themes they feel are relevant for the discussion on research culture and scientific integrity. From their answers, the facilitator, in consultation with the participants, prioritizes 6–9 themes. In addition, members of the research team take notes during the sessions to capture important elements.

Analysis

An inductive content analysis will be used to analyze 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 understudied topic). By using an inductive content analysis, we read through the data looking for recurring themes. First, the entire transcripts are read and emerging themes will be coded. New themes in the transcripts will be added to the list of codes and added to the previously analyzed results. The transcripts of the focus groups will be analyzed and coded independently by three team members (JT, TH, TBD). Individual analyses of the team members were compared and discussed to achieve consensus and to increase reliability.

To check validity, participants receive a written interpretation of the focus groups in which they participated, asking them to reflect on our interpretation and to indicate if they recognized the analysis and coding. This process of coding will yield the major themes (aim for approx. 8 themes). The results of the focus groups are then compared, analyzed and interpreted by the investigators, using an inductive approach. The final result is a summary of the themes. Typical quotes are identified per theme and per scientific rank (PhD student, postdoctoral fellow/assistant professor and associate professors/full professors) to clarify the coded themes. To optimize the quality of reporting, the COREQ checklist will be used (Tong et al, 2007)

2. When?

The focus group interviews will be performed in the second year of the project. We aim to start with the project January 2017. Consequently we hope to start this part of the project January 2018.

2. Preparation

List of necessities for the start of the qualitative part:

1. Semi structured focus group interview protocol (**Annex 5**), partially based on the results of the qualitative part
2. Audiotape electronics
3. Focus group interview facilities/rooms in the four academic institutions
4. Invitation of randomly selected participants
5. Training of TH to lead a focus group interviews/qualitative research methods

6.4 Pilot interventions

6.4.1. Who?

We will collaborate closely with two colleagues with relevant expertise from VU Medical Center (Prof. Guy A.M. Widdershoven (part 6.4.2.1A) and Prof. Yvo M. Smulders (part 6.4.2.1B), and Prof. Widdershoven will be closely involved in the research on moral case deliberation (part 6.4.2.1A). The part of the supervision program of novice PhD mentors will be developed during the first two years of the project (part 6.4.2.1B). We will gain information from experts in the field of academic education (i.e. the education department of the VU and VUmc) for designing an effective academic teaching module.

2. How?

The findings from the focus group interviews in combination with the relevant salient results of

the quantitative part will inform two pilot interventions to be tried out *during the third year of the project* within the academic community of the Vrije Universiteit Amsterdam. Participation in both these interventions will be voluntarily. The form of these interventions is predefined, but the content is flexible and will depend on the results of the first 2 years of the project.

6.4.2.1. Moral case deliberation

Moral awareness and an organizational climate that allows for moral learning are important preconditions for responsible conduct of research (RCR). Moral Case Deliberation (MCD) is an established intervention in health care that raises moral awareness and fosters an open culture (Van der Dam et al. 2015; Widdershoven & Metselaar, 2012). We will organize series of 5 MCDs within at least 4 research departments in different disciplinary fields on moral dilemmas related to research integrity that the researchers at issue experience themselves. The MCDs will be analyzed using thematic content analysis.

We aim to better understand the moral dilemmas that researchers encounter in their work, by uncovering the norms and values that underlie the dilemmas. Researchers are invited to participate in the MCDs via an invitational email that explains the purpose of MCDs. They are also asked to give informed consent after having had the time to ask any clarifying questions that arose from the invitational email. By participation in the MCDs, the researchers are supported when dealing with moral dilemmas and develop their moral competences. This is expected to affect the research climate within the participating departments, by stimulating the dialogue on dilemmas among researchers. Analyses of the MCDs will gain insight in the norms and values that evoke either RCR or QRP or worse.

6.4.2.2. Training program for novice PhD supervisors

Due to big fraud cases, increased awareness of research misbehavior (ranging from fraud to more subtle forms of research misbehavior) has resulted in research integrity training for PhD students as part of their educational program. However, training of the supervisors of these PhD students receives much less attention, whereas their good mentorship and role modelling are pivotal for responsible conduct of research (Bouter et al., 2016).

In this part of the project we will focus on the crucial role of supervisors in the education of PhD students. Novice PhD mentors will be invited to participate voluntarily, but the goal of the training is to set these mentors up with necessary competences to function as a successful mentor. With a training program for novice supervisors we will develop a module that consists of online activities, face-to-face education, regular one-on-one supervision and group meetings ('interview'). Our goal is to empower the novice supervisors for the important task of supervising PhD students with the ultimate goal to improve their skills to foster RCR in their PhD-students and in themselves.

We target specific competencies such as leadership, professionalism, collaboration, scholar skills and communication to increase research skills (van der Lee N, et al. Med Teach 2013;35:949-55.). The framework focusses on the development of the competences and skills necessary as a researcher and is primarily used to improve general competences, specific for researchers that work in a multidisciplinary environment. To evaluate this development and training we use the System of Evaluation of Teaching Qualities (SetQ) (Lombarts MJ, et al. Ned Tijdschr Geneesk 2010;154:A1222.). This tool aims to increase faculty's insight in teaching performance of mentors and trainees and is very suitable for our project as it gives us the opportunity to empirically evaluate our educational program to gain insight in the teaching performances of our trained mentors.

2. When?

Part 3a and 3b will be performed in the third year of the project. Start +/- November 2018.

2. Preparation

List of necessities for the start of part 3:

1. Moral case deliberation discussion leaders to guide the research groups bi-monthly with detailed guideline (**Annex 7**)
2. Detailed guideline for a module on fostering RCR for novice supervisors of PhD students (**Annex 8**)
3. Facilities (rooms, logistics, etc) to support the meetings and developing the teaching program.
4. 16 Novice supervisors of PhD students that are willing to participate in our pilot intervention
5. Intervention training program (written) + teaching modules and regular meetings for young/novice supervisors
6. Research groups of the 4 disciplinary fields that are willing to participate in our second intervention (moral case deliberation)

7. Relation of this protocol with the TWCF project

Project 1 of the TWCF proposal and this research protocol are complementary in that project 1 is *conceptual* and *normative* in nature, exploring the various epistemic values that the university ought to uphold, whereas this protocol is *empirical* in nature in that it investigates for specific epistemic values—those associated with academic integrity—to what extent they are actually pursued in research groups. It stands to reason that Responsible Conduct of Research (RCR) will involve intellectual virtues of various sorts. Therefore, there is further connection with the theoretical work done in project 3 of the TWCF which scrutinizes how groups can be intellectually virtuous. Since the humanities, including theology, are also among the academic disciplines from which research groups will be selected and to which the results of this project will be presented, this project will provide some information about how research groups in the humanities operate.

8. Envisioned output

The project will not only yield the usual output in the form of articles in scientific journals, but will also inform and inspire our academic communities and the leadership ??? the four institutions that collaborate in the Amsterdam Academic Alliance (AAA). The results of the web-based survey will be summarized for each faculty and division. The findings of the web-based survey and the focus group interviews will be discussed with the colleges of deans (VU and UvA) and the councils of division chairs (AMC and VUmc). We will make short reports per faculty/research institute to inform the deans of the faculties/institutes (only if $n > 25$). This will result in six reports for the faculties of the UvA and five reports for the faculties of the VU University.

We will also organize presentations for our academic communities about our findings, as well as panel discussions and debates about actions to be taken.

We will additionally present our findings on the national and international media and at academic conferences, like the 6th World Conference on Research Integrity.

Scientific output

- 6 peer-reviewed papers in journals such as *Science*, *Nature*, *PNAS*, *PLoS One*, *Research Integrity and Peer Review*, *J Empir Res Hum Res Ethics*, *Science and Engineering Ethics*, *American Scientist*, *Accountability in Research*
- 6 conference presentations at conferences such as the 6th *World Conference of Research Integrity*, the conference of the *International Center for Academic Integrity*, and the next version of the *REWARD Equator Conference* (<http://researchwaste.net/researchwasteequator->

conference/)

- 3 op-ed articles in leading Dutch newspapers, such as *Algemeen Dagblad*, *De Volkskrant*, *Trouw*, *NRC* as well as international news sources and blogs, such as *The Chronicle Higher Education* (60,000) and *The New Republic* (50,000)
- Detailed guideline for the use of moral case deliberation in research groups
- Detailed protocol for a module on fostering RCR for novice supervisors of PhD students

9. Project timeline

Qualitative part (Year 1)		
When?	What?	Who?
1-3 months	Writing research protocol (including approval of rectors and deans offices of the 4 institutions)	JT, TH, LB
1-3 months	METC approval at VUmc and UvA	TH, JT and LB
1-3 months	Designing PhD training program	TH, JT, LB
3-6 months	Collecting and checking emailaddresses + other practical preparations from 4 institutions	TH, JT
1-6 months	Preparing and piloting Survey	TH, JT, LB, BM
7-9 months	Sending survey + 2 reminders	TH, JT
10-12 months	Data analysis + reporting to prepare focus groups for project 2	TH, JT, LB
13-24 months	Writing first manuscript	TH, JT, LB
1-12 months	Continuous trainingprogram for TH student Training qualitative research methods	TH
Month 12	Evaluation of progress of TH	TH, JT, LB
Year 2, part 2		
1-3 months	Preparing focus group interviews	TH, JT, to be determined
1-3 months	Writing focus group protocol	TH, JT, LB
1-3 months	Preparing practicalities (invitations, location, planning) for focus group interviews	TH
1-3 months	Performing pilot focus group and writing semistructured protocol	TH, JT
4-9 months	Conducting focus group interviews	TH, JT
7-9 months	Transcribing interviews	TH
10-18 months	Writing manuscript	TH, JT, LB
10-12 months	Preparing intervention project + writing moral case deliberation protocol	TH, JT, LB, GW
10-12 months	Writing teaching module /protocol for novice supervisors	TH, JT, LB, (evt. External help)
Year 3: Part 3a and b, intervention studies		

1-3 months	Completing moral case deliberation (MCD) protocol (Annex 7)	TH, JT, LB, GW
1-3 months	Completing novice supervisors protocol (Annex 8)	TH, JT, LB, (external)
1-3 months	Organizing/planning intervention program with MCD and intervention groups	TH, JT, LB
1-3 months	Organizing supervisors training protocol	TH, JT, LB
4-9 months	Conducting MCD and supervisors protocol	TH, JT, LB
10-12 months	Analyzing results of intervention study (see Annex 7 for complete description of analysis and expected results)	TH, JT, LB
10-12 months	Writing 2 manuscripts of intervention studies	TH, JT, LB
10-12 months	Completing reports, conducting process evaluation report with recommendations	TH, JT, LB
Year 4: Writing PhD thesis		
1-12 months	Extension time for publications (see above)	TH, JT, LB
1-12 months	Extended time to finalize quantitative, qualitative and intervention part.	TH, JT, LB
1-12 months	Writing dissertation	TH, JT and LB
1-12 months	Writing 2 papers with the other team members of other projects of the TWCF (Annex 1) to merge and combine the results of this project with the other projects.	JT, LB, RvW, RP, JdR

10. Budget

— PhD student (1.0 fte for 4 years)

The 4th year is for finishing the PhD thesis and the manuscripts of the articles based on the AAA project. Furthermore, in this last year a systematic review or a psychometric study on one of the three instruments used will be performed.

— Postdoc (0.4 fte for 4 years)

— Other costs (75 K€)2

PhD training program (12K€), Conference meetings + travel costs postdocs and PhD (3 years x 6k€ = 18k€), chairs and data-analysts focus group interviews (20 K€), external location and catering for focus group interviews (10 K€), and travel budget for postdoc and US partner (Brian Martinson) (15 K€)

References

1. Bouter LM. Perverse incentives or rotten apples? Accountability in Research 2015; 22: 148–61
2. Martinson BC, Thrush CR, & Crain AL. 2013. Development and validation of the Survey of Organizational Research Climate. Sci Eng Ethics 2013; 19: 813–34
3. Crain AL, Martinson BC, Thrush CR. Relationships between the Survey of Organizational Research Climate and self-reported research practices. Sci Eng Ethics 2013; 19: 835-40

4. Wells JA, Thrush CR, Martinson, BC, May, TA, Stickler, M, Callahan, EC, Klomparens KL 2014. Survey of organizational research climate. *J Emp Res Hum Res Ethics* 2015; 9: 72–88
5. Tijdink JK, Smulders YM, Vergouwen AC, de Vet HC, Knol DL 2014. The assessment of publication pressure in medical science: validity and reliability of a Publication Pressure Questionnaire (PPQ). *Qual Life Res* 2014; 23: 2055–62
6. Tijdink JK, Vergouwen AC, Smulders YM. 2013. Publication pressure and burn out among Dutch medical professors: a nationwide survey. *PLoS One* 2013; 8: e73381
7. Tijdink JK, Verbeke R, Smulders YM 2014. Publication pressure and scientific misconduct in medical scientists. *J Emp Res Hum Res Ethics* 2014; 9: 64–71
8. Creswell JW. *Qualitative inquiry and research design: choosing among five traditions*. Thousand Oaks: Sage; 1998
9. Krueger RA, Casey MA. *Focus groups: a practical guide for applied research*. 3rd ed. Thousand Oaks CA: Sage Publications; 2000
10. Tijdink JK, Schipper K, Bouter LM, Maclaine Pont P, de Jonge J, Smulders YM. Biomedical scientists' view on publication culture: a focus group study. *BMJ Open* 2016;6:e008681
11. Mays N, Pope C. *Qualitative research in health care. Assessing quality in qualitative research*. *BMJ* 2000; 320: 50-2
12. Meadows LM, Morse JM. Constructing evidence within the qualitative project. In: Morse JM, Swanson JM, Kuzel AJ, editors. *The Nature of qualitative evidence*. Thousand Oaks, CA: Sage publications; 2001. p. 187-200
13. Tong D, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*. 2007. Volume 19, Number 6: pp. 349 – 357)
14. Bouter, L. M., Tijdink, J., Axelsen, N., Martinson, B. C., & ter Riet, G. (2016). 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(1). doi:10.1186/s41073-016-0024-5

Appendix 3. List of 60 misbehaviours in survey per category and in no particular order, taken from (1).

Study design
<ol style="list-style-type: none"> 1. Propose study questions which are clearly irrelevant 2. Choose a clearly inadequate research design or using evidently unsuitable measurement instruments 3. Present grossly misleading information in a grant application 4. Write no or a clearly inadequate research protocol 5. Ignore substantial safety risks of the study to participants, workers or environment 6. Ignore substantial risks of the expected findings for society or environment 7. Importantly change the research design during the study without disclosure 8. Give insufficient attention to the equipment, skills or expertise which are essential to perform the study
Data collection
<ol style="list-style-type: none"> 9. Collect more data after noticing that the results are almost statistically significant 10. Fabricate data 11. Stop data collection earlier than planned because the results are already statistically significant 12. Not adhere to pertinent laws and regulations 13. Inadequately handle or store data or materials 14. Keep inadequate notes of the research process 15. Ignore basic principles of quality assurance
Reporting
<ol style="list-style-type: none"> 16. Report on data-driven hypotheses without disclosure 17. Delete data before performing data analysis without disclosure 18. Selectively delete data, modify data or add fabricated data after performing initial data-analyses 19. Perform data-analyses not stated in the study protocol without disclosure 20. Report an incorrect downwardly rounded p-value 21. Not report all study protocol-stipulated results 22. Not publish a valid 'negative' study 23. Report an unexpected finding as having been hypothesized from the start 24. Conceal results that contradict earlier findings or convictions

25. Not report clearly relevant details of study methods
26. Not report replication problems
27. Selectively cite to enhance own findings or convictions
28. Selectively cite to please editors, reviewers or colleagues
29. Selectively cite or cite own work to improve citation metrics
30. Let own convictions influence the conclusions substantially
31. Insufficiently report study flaws and limitations
32. Spread study results over more papers than needed
33. Duplicate publication without disclosure
34. Re-use parts of own publications without referencing
35. Re-use of previously published data without disclosure
36. Modify the results or conclusions of a study due to pressure of a sponsor
37. Failure to disclose a sponsor of the study
38. Failure to disclose a relevant financial or intellectual conflict of interest
39. Handle existing conflicts of interest inadequately
40. Communicate results to the general public before a peer reviewed publication is available
41. Deliberately communicate findings inaccurately in the media or during presentations
42. Make no clear distinction between personal views and professional comments

Collaboration

43. Take no full responsibility for the integrity of the research project and its reports
44. Refuse to share data with bona fide colleagues
45. Turn a blind eye to putative breaches of research integrity by others
46. Refuse to respond to an allegation of a breach of research integrity
47. Use unpublished ideas or phrases of others without their permission
48. Use published ideas or phrases of others without referencing
49. Unfairly review papers, grant applications or colleagues applying for promotion
50. Review one's own submitted manuscripts
51. Demand, accept or offer substantial gifts for doing a favour
52. Insufficiently supervise or mentor junior coworkers
53. Gross unfairness to collaborators
54. Add an author who doesn't qualify for authorship
55. Demand or accept an authorship without significant contribution

56. Omit a contributor who deserves authorship
 57. Not acknowledge contributors who do not qualify for authorship
 58. Not ask permission from contributors for the wording of the acknowledgement
 59. Not share reviewers' comments with all co-authors
 60. Submit or resubmit a paper or grant application without consent from all authors
-

Appendix 4. Focus groups protocol

Introduction (10 min)

- Thank the focus group participants for agreeing to the focus group
- Introduce focus group leader and observant as well as participants
- Check informed consent, if it is ok to record the interview, privacy policy and anonymity
- Explain what will happen with the results
- Explanation of how a focus group works, what are the rules for group discussions, etc.
- Ask participants if they have any questions before the focus group begins
- Brief introduction of this research project (refresh)

Interview topics (3) to be covered during the focus group (80 min)

1. Interactive assignment (think, pair, share). Assessment of the question: “What do you envision when thinking of a responsible research climate?” (25 min)

Asks participants to individually write down three characteristics of responsible research climate (5 min). Participants are then invited to exchange these with their neighbours (5 min). Finally, groups are asked to share their characteristics with the group and discuss them deeper (10- 15 min)

2. Interactive assignment using post-its to identify misbehaviours (25 min). Participants are presented with specific misbehaviours we found important in their field and consequently invited to think of misbehaviours in their own research climate in their specific field and write down about 3-5 ones that play a big role. Small groups present their misbehaviours. (10 min)

Verification of our findings that regard the research climate for this specific group (field X with rank Y) by means of providing them with the 5 misbehaviours. They are then asked to rank these behaviours by means of severity as a group using a major-minor scale. (5 min)

The group as a whole is asked order these items on the major-minor scale including the new misbehaviours (10 min).

Optional short break (10 min)

3. Group discussion: “Which barriers do you perceive in the research climate for responsible conduct of research?” (20 min)

4. How would you tackle these barriers? (10 min)

I (see above)

II

II etc.

Ending the interview (10 min)

- Is there anything else you'd like to say?
- What is the main message for us to take away?
- Can we contact you in case we need any additional information or if something is unclear?
- What would you like to see coming out of our study?
- Report of this meeting with the option to comment and suggest corrections and mention the summary that will be sent to all participants to comment (member-check)
- Close the interview and thank the interviewees for their participation.

Clustering themes are dark blue and the overall theme is light green. Lines signify relations (i.e. is one example of/is part of/is cause of).

Appendix 6. Top 5 most frequent research misbehaviours per academic rank and disciplinary field.

Top 5														
Academic rank, <i>M (SD)</i>							Disciplinary field, <i>M (SD)</i>							
PhD students			Postdocs & assistant professors		Associate & full professors		Biomedicine		Natural sciences		Social sciences		Humanities	
#1	Insufficiently mentor or supervise junior co-workers	2.03 (.77)	Demand or accept an authorship without significant contribution	2.15 (.76)	Selectively cite to please editors, reviewers or colleagues	2.19 (.70)	Add an author who doesn't qualify for authorship	2.14 (.71)	Insufficiently supervise or mentor junior co-workers	2.19 (.78)	Insufficiently supervise or mentor junior co-workers	2.04 (.82)	Selectively cite to please editors, reviewers or colleagues	2.03 (.81)
#2	Demand or accept an authorship without significant contribution	2.03 (.77)	Insufficiently supervise or mentor junior co-workers	2.1 (.75)	Selectively cite to enhance own findings or convictions	2.09 (.70)	Demand or accept an authorship without significant contribution	2.14 (.74)	Selectively cite to please editors, reviewers or colleagues	2.16 (.71)	Demand or accept an authorship without significant contribution	2 (.78)	Let own convictions influence the conclusions substantially	2.03 (.90)
#3	Add an author who doesn't qualify for authorship	1.97 (.73)	Selectively cite to please editors, reviewers or colleagues	2.09 (.77)	Demand or accept an authorship without significant contribution	2.09 (.72)	Insufficiently supervise or mentor junior co-workers	2.06 (.72)	Not report clearly relevant details of study methods	1.98 (.70)	Selectively cite to enhance own findings or convictions	1.96 (.75)	Demand or accept an authorship without significant contribution	2.03 (.74)
#4	Keep inadequate notes of the research process	1.95 (.73)	Selectively cite or cite own work to improve citation metrics	2.04 (.75)	Selectively cite or cite own work to improve citation metrics	2.07 (.77)	Keep inadequate notes of the research process	1.93 (.71)	Keep inadequate notes of the research process	1.95 (.78)	Selectively cite to please editors, reviewers or colleagues	1.95 (.84)	Insufficiently supervise or mentor junior co-workers	2.02 (.77)
#5	Selectively cite to please editors, reviewers or colleagues	1.78 (.75)	Add an author who doesn't qualify for authorship	1.98 (.78)	Insufficiently supervise or mentor junior co-workers	2.04 (.75)	Selectively cite to please editors, reviewers or colleagues	1.88 (.74)	Demand or accept an authorship without significant contribution	1.95 (.84)	Communicate results to the general public before a peer reviewed publication is available	1.88 (.81)	Selectively cite to enhance own findings or convictions	2.00 (.74)

M = mean score per subgroup, *SD* = standard deviation. Frequency was scored on a 0 = 'never, 1 = 'once or twice' and 3 = three times or more scale. The higher the mean score, the more frequent the misbehaviour was perceived.

Appendix 7. Top 5 most impactful research misbehaviours per academic rank and disciplinary field.

Top 5														
Academic rank, <i>M (SD)</i>							Disciplinary field, <i>M (SD)</i>							
PhD students			Postdocs & assistant professors		Associate & full professors		Biomedicine		Natural sciences		Social sciences		Humanities	
#1	Fabricate data	4.55 (1.10)	Fabricate data	4.70 (.84)	Fabricate data	4.23 (1.3)	Fabricate data	4.51 (1.09)	Fabricate data	4.53 (1.28)	Fabricate data	4.62 (.94)	Fabricate data	4.44 (1.16)
#2	Modify the results or conclusions of a study due to pressure of a sponsor	4.41 (.94)	Selectively delete data, modify data or add fabricated data after performing initial data-analyses	4.4 (.96)	Selectively delete data, modify data or add fabricated data after performing initial data-analyses	4.22 (1.21)	Modify the results or conclusions of a study due to pressure of a sponsor	4.15 (1.16)	Modify the results or conclusions of a study due to pressure of a sponsor	4.33 (1.06)	Selectively delete data, modify data or add fabricated data after performing initial data-analyses	4.49 (.93)	Modify the results or conclusions of a study due to pressure of a sponsor	4.15 (1.11)
#3	Selectively delete data, modify data or add fabricated data after performing initial data-analyses	4.05 (1.32)	Modify the results or conclusions of a study due to pressure of a sponsor	4.22 (1)	Conceal results that contradict earlier findings or convictions	4.03 (.89)	Selectively delete data, modify data or add fabricated data after performing initial data-analyses	4.13 (1.21)	Selectively delete data, modify data or add fabricated data after performing initial data-analyses	4.17 (1.28)	Modify the results or conclusions of a study due to pressure of a sponsor	4.48 (.83)	Present grossly misleading information in a grant application	3.94 (.89)
#4	Conceal results that contradict earlier findings or convictions	3.87 (1.06)	Review one's own submitted manuscript	4.09 (1.08)	Choose a clearly inadequate research design or using evidently unsuitable measurement instruments	3.9 (1)	Review one's own submitted manuscript	3.92 (1.17)	Refuse to respond to an allegation of a breach of research integrity	3.92 (.96)	Modify the results or conclusions of a study due to pressure of a sponsor	3.99 (.96)	Selectively delete data, modify data or add fabricated data after performing initial data-analyses	3.94 (1.35)
#5	Choose a clearly inadequate research design or using	3.85 (1.18)	Choose a clearly inadequate research design or using	3.88 (.96)	Modify the results or conclusions of a study due to	3.88 (1.46)	Choose a clearly inadequate research design or using	3.91 (1.01)	Review one's own submitted manuscript	3.8 (1.46)	Review one's own submitted manuscripts	3.95 (1.28)	Choose a clearly inadequate research design or using	3.81 (1.13)

evidently unsuitable measurement instruments	evidently unsuitable measure- ment instruments	pressure of a sponsor	evidently unsuitable measure- ment instruments	evidently unsuitable measure- ment instruments
---	--	--------------------------	--	--

M = mean score per subgroup, SD = standard deviation. Impact was scored on a 1 = negligible and 5 = major impact scale. The higher the mean score, the more impact the misbehaviour was perceived to have on the validity of the study's findings.

Appendix 8. Top 5 most important (defined as aggregated frequency * impact) research misbehaviours per academic rank.

Top 5						
Academic rank, <i>M (SD)</i>						
PhD students			Postdocs & assistant professors		Associate & full professors	
#1	Insufficiently mentor or supervise junior coworkers	7.05 (3.76)	Insufficiently supervise or mentor junior coworkers	7.55 (3.75)	Let own convictions influence the conclusions substantially	6.74 (3.29)
#2	Give insufficient attention to the equipment, skills or expertise which are essential to perform the study	5.9 (3.41)	Choose a clearly inadequate research design or using evidently unsuitable measurement instruments	6.23 (3.1)	Insufficiently supervise or mentor junior coworkers	6.37 (3.65)
#3	Let own convictions influence the conclusions substantially	5.86 (3.17)	Let own convictions influence the conclusions substantially	6.11 (3.18)	Choose a clearly inadequate research design or using evidently unsuitable measurement instruments	6 (2.85)
#4	Choose a clearly inadequate research design or using evidently unsuitable measurement instruments	5.7 (3.44)	Not publish a valid 'negative' study	6 (3.62)	Not publish a valid 'negative' study	5.95 (3.57)
#5	Inadequately handle or store data or materials	5.64 (3.13)	Insufficiently report study flaws and limitations	5.83 (3.36)	Insufficiently report study flaws and limitations	5.94 (2.75)

M = mean score per subgroup, *SD* = standard deviation. Importance was computed as the product score of frequency (1-3) and impact (1-5) and thus ranged from 1 to 15. The higher the mean score, the more important this research misbehaviour was perceived according to researchers from a particular rank.

Appendix 9. 60 items ranked for importance per disciplinary field and academic rank.

<i>Items written out</i>	Total sample			PhD students			PD & Asis.Prof			Asso & Prof			Biomed			Nat			Soc			Hum		
	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>	<i>N</i>	<i>M</i>	<i>SD</i>
Propose study questions which are clearly irrelevant	356	4.3	2.9	158	3.9	2.5	107	4.3	2.8	80	4.8	3.5	199	4.1	2.6	43	3.9	2.3	79	4.5	3.4	33	4.6	3.5
Choose a clearly inadequate research design or using evidently unsuitable measurement instruments	348	5.9	3.2	166	5.7	3.4	101	6.2	3.1	67	6	2.8	192	6	3.2	40	5.5	3.2	78	5.8	3.4	37	6.1	3.4
Present grossly misleading information in a grant application	357	4.1	1.8	162	4	1.9	101	4.1	1.5	72	4.5	1.9	209	4.1	1.7	43	5	2.7	70	3.7	1.6	31	4.3	1
Collect more data after noticing that the results are almost statistically significant	351	4.3	2.4	158	4.2	2.3	112	4.4	2.7	68	4.4	2.3	195	4.3	2.3	37	4.3	2.7	79	4.4	2.6	39	3.8	2.2
Fabricate data	355	4.8	1.7	172	4.8	1.7	105	4.9	1.2	70	4.8	2.2	199	4.7	1.4	36	5.5	2.8	84	4.8	1.4	34	4.9	2.1
Report on data-driven hypotheses without disclosure	349	4.2	2.4	162	4	2.2	105	4.5	2.5	72	4.5	2.6	193	4.2	2.4	40	3.6	1.8	81	4.9	2.8	34	3.9	1.3
Delete data before performing data analysis without disclosure	355	4.3	2.1	166	4.3	2.1	101	4.4	2	77	4.2	2.2	198	4.3	1.8	33	4.5	3.5	79	4.5	1.9	42	3.5	1.9
Selectively delete data, modify data or add fabricated data after performing initial data-analyses	348	4.8	2	153	4.7	2.1	108	5.1	2	73	4.7	2	185	4.7	1.8	41	5.4	3.3	85	5	1.6	35	4.5	1.7
Perform data-analyses not stated in the study protocol without disclosure	358	4.5	2.6	166	4.5	2.7	106	4.8	2.8	70	4.2	2.3	199	4.5	2.6	37	3.6	2	83	5	2.8	36	3.8	2.6
Report an incorrect downwardly rounded p-value	361	3.5	1.7	169	3.6	1.6	105	3.6	1.8	72	3.5	1.8	210	3.6	1.7	39	3.4	1.6	77	3.7	1.6	33	3	1.6
Not report all study protocol-stipulated results	347	4.5	2.6	150	4.3	2.6	108	4.7	2.7	73	4.4	2.5	185	4.3	2.1	33	4.4	3.2	85	5	3.4	41	4	2
Not publish a valid 'negative' study	355	5.6	3.4	159	5.5	3.4	117	6	3.6	63	5.5	3.2	198	5.5	3.2	40	5.5	3.7	86	6.5	4	30	4.2	2.4
Report an unexpected finding as having been hypothesized from the start	357	4.7	2.8	173	4.7	2.8	100	5	2.8	67	4.6	2.7	190	4.6	2.5	39	4	2.9	87	5.6	3.1	39	4.2	2.9
Conceal results that contradict earlier findings or convictions	344	5.2	2.5	156	5.2	2.5	101	4.9	2.5	75	5.5	2.4	181	5.2	2.4	41	5	2.8	82	5.3	2.7	38	4.8	2.2
Take no full responsibility for the integrity of the research project and its reports	346	4	2.3	168	4	2.3	107	4	2.1	58	4.3	2.8	200	3.9	2.2	34	5.1	3	72	4.1	2.4	37	3.7	1.8
Refuse to share data with bona fide colleagues	354	4.2	2.7	165	4	2.5	102	4.7	3.1	79	4	2.3	190	4.3	2.5	42	4.5	3.5	87	4.1	2.6	34	3.6	2.3
Turn a blind eye to putative breaches of research integrity by others	352	4.5	2.4	164	4.3	2	94	5	3.1	75	4.6	2.5	203	4.4	2.2	35	4.5	2.4	85	4.8	2.8	27	4.7	2.9
Refuse to respond to an allegation of a breach of research integrity	356	3.9	1.5	170	3.7	1.4	106	4	1.5	66	4.1	1.9	194	3.6	1.5	39	4.2	1.6	87	4.2	1.6	34	3.9	1.5
Use unpublished ideas or phrases of others without their permission	352	3.5	2.2	160	3.4	2	103	3.8	2.3	75	3.5	2.6	200	3.5	2.1	35	3.5	1.9	83	3.4	2	32	4.2	3.5
Use published ideas or phrases of others without referencing	356	3.8	2.3	164	3.5	2	109	4.1	2.6	69	4.2	2.5	206	3.6	2	34	3.9	3	87	3.9	2.5	28	5.1	3
Write no or a clearly inadequate research protocol	353	4.9	2.8	159	4.9	2.8	104	4.8	2.5	80	5.3	3	208	5.2	2.7	44	5.3	3.3	70	4.4	2.6	29	4.1	2.3
Ignore substantial safety risks of the study to participants, workers or environment	360	3.6	1.8	175	3.7	1.7	102	3.7	1.9	68	3.2	2	210	3.7	1.8	42	3.5	1.7	74	3.6	1.7	33	2.8	1.9
Stop data collection earlier than planned because the results are already statistically significant	350	3.8	2	162	3.8	1.6	108	3.9	2.3	70	3.7	2.3	189	3.8	1.7	43	3.4	2.3	76	4	2.2	42	3.9	2.4
Not adhere to pertinent laws and regulations	350	3.8	1.7	170	3.8	1.7	87	3.8	1.6	79	3.7	1.9	184	3.8	1.9	37	3.8	1.9	91	3.5	1.5	37	4.2	1.6
Not report clearly relevant details of study methods	356	5.2	3	169	4.9	2.8	109	5.7	3.3	71	5.3	3.1	205	5.1	2.9	40	7	3.4	78	5.3	3	32	3.4	1.8
Not report replication problems	353	4.7	2.7	155	4.7	2.9	109	5.1	2.3	69	4.1	2.7	187	4.6	2.4	49	4.3	2.8	88	5.2	2.9	29	4	3
Selectively cite to enhance own findings or convictions	359	5.5	2.9	157	5.1	2.9	115	5.8	2.7	77	5.9	2.7	210	5.5	2.9	45	5.2	2.7	73	5.5	2.4	30	6.2	3.2
Selectively cite to please editors, reviewers or colleagues	344	4.8	2.8	161	4.5	2.7	108	5.2	2.8	62	4.9	2.9	192	4.4	2.4	45	5	3.2	65	5	3.1	39	6	3.1
Selectively cite or cite own work to improve citation metrics	344	4.2	2.5	177	3.9	2.4	100	4.6	2.7	54	4.4	2.8	203	4.2	2.5	34	3.9	2.8	73	4.2	2.2	32	4.1	3.3
Let own convictions influence the conclusions substantially	362	6.1	3.2	171	5.9	3.2	104	6.1	3.2	69	6.7	3.3	202	6	3.2	37	6.4	3.3	84	5.9	2.9	35	6.7	3.7
Insufficiently report study flaws and limitations	341	5.5	3.1	156	5.4	2.9	101	5.8	3.4	70	5.6	3.1	196	5.4	3	36	6.6	3.4	71	5.2	3.3	36	5.4	2.6
Spread study results over more papers than needed	353	4.1	2.6	160	3.9	2.6	104	4.5	2.7	72	4.1	2.4	205	4.1	2.6	36	4.1	2.3	74	4.3	2.7	34	4.1	2.7
Duplicate publication without disclosure	362	3.3	1.7	156	3.4	1.7	115	3.3	1.8	73	3.1	1.7	191	3.2	1.5	44	3.9	2.3	87	3.3	1.6	36	3.2	1.7
Re-use parts of own publications without referencing	354	3.2	1.8	173	2.9	1.7	94	3.5	1.9	76	3.7	2	213	3	1.8	37	3.1	1.5	70	3.2	1.8	33	4.2	2.4
Unfairly review papers, grant applications or colleagues applying for promotion	357	4.9	3	169	4.1	2.2	101	5.7	3.6	73	5.9	3.6	208	4.6	2.7	45	5.3	3.6	67	5.3	3	34	6	4.1
Review one's own submitted manuscripts	364	4.1	1.5	161	3.9	1.3	111	4.4	1.6	78	4.1	1.6	207	4.1	1.4	45	3.9	1.4	79	4.1	1.6	30	3.9	1.5
Demand, accept or offer substantial gifts for doing a favour	357	3.4	1.6	172	3.3	1.5	99	3.5	2	71	3.4	1.4	203	3.2	1.5	40	4.2	2.4	76	3.2	1.3	36	3.8	1.4
Insufficiently supervise or mentor junior coworkers	355	7.1	3.7	157	7.1	3.8	107	7.6	3.8	78	6.4	3.7	197	7	3.6	32	7.7	4.1	77	6.9	3.8	46	6.8	3.8
Gross unfairness to collaborators	348	3.9	2.6	145	3.7	2.4	106	4	2.8	81	4	2.3	207	3.9	2.4	34	4.5	2.8	75	3.7	2.5	30	4	3.3
Submit or resubmit a paper or grant application without consent from all authors	364	3.2	1.9	172	3.1	1.8	118	3.3	2.1	59	3.3	2.1	204	3.3	2	39	3.2	1.7	81	3	2	38	2.9	1.6
Ignore substantial risks of the expected findings for society or environment	340	3.7	1.8	160	3.8	1.6	95	3.6	1.6	70	3.9	2.6	197	3.8	1.5	37	3.4	1.9	69	3.7	2	33	3.5	3
Importantly change the research design during the study without disclosure	354	4.4	2.1	167	4.3	2.1	103	4.7	2.3	67	4.2	2	190	4.4	2.1	32	4.3	2.5	92	4.3	2.2	39	4.4	2.2
Give insufficient attention to the equipment, skills or expertise which are essential to perform study	344	5.6	3.4	162	5.9	3.4	110	5.6	3.2	55	5.1	3.5	183	5.6	3.3	39	6.3	3.5	89	5.7	3.3	30	3.9	2.9
Inadequately handle or store data or materials	356	4.6	2.8	164	4.9	3	102	4.3	2.6	76	4.6	2.8	207	4.9	2.8	38	4.8	3.2	72	4.3	2.9	38	3.7	2.5
Keep inadequate notes of the research process	357	5.5	3.1	167	5.6	3.1	105	5.6	3.2	73	5.2	2.8	200	5.6	3	40	6	3.7	81	5.4	3	34	4.5	3
Ignore basic principles of quality assurance	356	5.3	2.8	170	5	2.8	103	5.5	2.6	79	5.8	3.1	202	5.2	2.8	41	5.3	2.8	80	5.4	2.9	33	5.3	2.5
Re-use of previously published data without disclosure	352	3.2	1.7	174	3.1	1.4	84	3.5	2	83	3.3	2.1	211	3.1	1.5	37	3	1.6	71	3.5	2.3	32	3.6	1.7
Modify the results or conclusions of a study due to pressure of a sponsor	361	4.6	1.5	172	4.8	1.4	117	4.5	1.2	59	4.3	2.3	194	4.5	1.5	43	4.7	1.7	85	4.8	1.6	39	4.4	1.4
Failure to disclose a sponsor of the study	353	3.2	1.5	171	3.1	1.4	111	3.2	1.5	59	3.4	1.8	191	3.2	1.5	40	3.1	1.4	81	3.3	1.6	38	3.1	1.3
Failure to disclose a relevant financial or intellectual conflict of interest	358	3.6	1.8	167	3.4	1.5	97	3.6	1.9	78	3.8	2.4	197	3.6	1.8	35	3.2	1.9	81	4	2	45	4.7	2.9
Handle existing conflicts of interest inadequately	346	4.1	2.5	165	3.8	2.2	97	4.5	2.5	71	4.4	3	192	4.1	2.4	40	4.2	2.8	68	3.8	2.4	33	3.8	1.4
Communicate results to the general public before a peer reviewed publication is available	358	3.9	2.2	172	3.7	2	105	3.9	2.4	64	4.4	2.4	194	3.9	2.1	39	3.5	2	90	4.1	2.5	37	5.2	3.5
Deliberately communicate findings inaccurately in the media or during presentations	368	4.6	2.5	163	4.5	2.6	118	4.7	2.5	71	4.8	2.7	193	4.5	2.3	46	4.5	2.7	87	4.6	2.5	38	5.7	3.7
Make no clear distinction between personal views and professional comments	338	4.9	3.1	157	4.4	2.9	102	5.1	3.3	68	5.4	3.1	176	4.7	2.8	37	4.3	2.9	85	5.2	3.4	43	3.7	3.4
Add an author who doesn't qualify for authorship	355	3.7	2.5	179	3.9	2.6	93	4	2.6	66	3.2	1.9	207	3.9	2.3	36	3.7	2.8	67	3.4	1.9	32	5.2	3.5
Demand or accept an authorship without significant contribution	346	4.3	2.8	162	4.3	2.9	95	4.7	3.1	78	3.9	2.3	197	4.3	2.8	43	3.6	2.6	73	4.2	2.8	38	3.7	2.2
Omit a contributor who deserves authorship	352	3.5	2.3	166	3.4																			

Appendix 10. Qualitative rankings of important research misbehaviours per focus group.

Important research misbehaviours stratified per academic rank and disciplinary field			
Biomedical sciences researchers	Natural sciences researchers	Social sciences researchers	Humanities researchers
PhD students	PhD students	PhD students	PhD students
Bad administration/no raw data	No mechanism for reporting errors in already published papers	<i>Let own convictions influence conclusions substantially</i>	Discrimination
Inadequate power analysis	Hiding important information from the data with wrong statistical test (intentionally or unintentionally)	An emphasis on what is known instead of stimulating the creative process of thinking outside the box	Manipulation by stakeholder
<i>Choose a clearly inadequate research design or using evidently unsuitable measurement instruments</i>	<i>Insufficiently supervising or mentoring junior co-workers</i>	Publication pressure preventing proper skill- and knowledge development by putting timeframes on writing dissertation & writing & reviewing papers	Intentional destruction of artefacts
Selectively choosing primary outcome study	Changing the parameters in the method, not to allow the other scientists reproduce the work quickly	Time management, publication and teaching obligation spiral in the tenure track	Guest authorship

<i>Keep inadequate notes of the research process</i>	Withholding material for reproduction of published work	Peer reviewing seems overly critical, we are trained to see what is “bad” instead of seeing what is “good”. Especially in small fields with same reviewers for different journals, a reviewer has extensive power	<i>Unfairly review papers, grants, colleagues applying for promotion</i>
--	---	--	--

Postdocs and assistant professors	Postdocs and assistant professors	Postdocs and assistant professors	Postdocs and assistant professors
<i>Let own convictions influence conclusions</i>	Strengthening your own position through power abuse	Bullying juniors	Stealing someone else’s PhD (allowing them to finish and defend)
Considering an opinion as more valuable than objective indicators of good research	A senior research that gets too many PhD candidates and postdocs; group relations may sour and leads to the PhD candidates not getting adequate supervision	Pressuring junior postdocs and PhDs in “finding” significant effects	Taking PhD students work for publication without acknowledgement/ Using students’ work for publication with little or no acknowledgement
Investigating the null hypothesis less thoroughly than the alternative hypothesis	Inadequate supervision (linked to following items)	Pressure to work on grants while you are still working on your former project (PhD project)	Submitting publishing the same material (text) in different journals chapters

Lack of doubt	Writing a manuscript chronologically	Adjust covariates to get good results/ Adjust correction for multiple testing	Sending PhD students here for supervision (from abroad, since actual supervisor is not an expert)
Getting too little time for research from mentor	Failing to attribute own success to team effort	Supervisors exploiting PhDs/postdocs via their power positions (co-authorships/ harassments)	<i>Insufficiently supervise or mentor junior co-workers</i>
Associate and full professors	Associate and full professors	Associate and full professors	Associate and full professors
Wrong role models	Falsifying results	Not carrying out replication study/ not reporting of non-replication	Plagiarism
<i>Insufficiently supervise or mentor junior co-workers</i>	Plagiarism	<i>Choose a clearly inadequate research design or evidently unsuitable measurement instruments</i>	Reviewer rejecting paper and then stealing ideas
Not being transparent in reporting/ keep inadequate notes of research process	Declining or withholding publication by reviewing competitor	Report an unexpected finding as being hypothesised from the start	Data manipulation (adjustment analyses)
Not publishing results	Abuse of referee or editor power/ Referees that steal	Selective data reporting	Making up data

	ideas from proposals/manuscripts they need to evaluate		
Leaving out data/Selectively including/excluding data	Espionage/ stealing (conscious or unconscious)	Publication bias	Ideologizing science / Economising science / Politicalising science

* *Italics* denote research misbehaviours that came from the quantitative survey. Focus groups were conducted over 3 academic ranks (see rows, PhD students, postdocs and assistant professors and associate and full professors) and 4 academic ranks (see columns, biomedical sciences, natural sciences, social sciences and humanities). The higher the misbehaviour in the rows; the more important focus group participants indicated this research misbehaviour to be.

Appendix 11. List of 60 misbehaviours with alternative formulations and newly added items from the focus group interviews printed in *italics*. Newly added and preferred formulations of existing items are printed ***italic-bold*** and the former formulation is ~~crossed-out~~.

Study design

1. Propose study questions which are clearly irrelevant
2. Choose a clearly inadequate research design or using evidently unsuitable measurement instruments
3. Present grossly misleading information in a grant application
Overpromising future research results in a grant application (knowing it's impossible to obtain these results)
4. Write no or a clearly inadequate research protocol
5. Ignore substantial safety risks of the study to participants, workers or environment
6. Ignore substantial risks of the expected findings for society or environment
7. Importantly change the research design during the study without disclosure
8. Give insufficient attention to the equipment, skills or expertise which are essential to perform the study
Not keeping up with scientific developments in your field

Data collection

9. Collect more data after noticing that the results are almost statistically significant
10. Fabricate data
Make up data
11. Stop data collection earlier than planned because the results are already statistically significant
12. Not adhere to pertinent laws and regulations
13. Inadequately handle or store data or materials

Un-anonymised data storage without consent

14. Keep inadequate notes of the research process
15. Ignore basic principles of quality assurance
16. *Publish datasets unorganised in an Open Access database that are not usable for interested researchers*
17. ***Violating medical ethical guidelines or privacy guidelines***
18. *Inadequate or no power analysis*

To not conduct a power calculation or sample size calculation, although required

Insufficient attention to statistical power in the research design

Change intended sample size after data collection without disclosure

Reporting

19. Report on data-driven hypotheses without disclosure

Hedging of chances by including more variables (dependent or independent) than the primary study entails in their research protocol

Changing the research question after results are known
20. Delete data before performing data analysis without disclosure
21. Selectively delete data, modify data or add fabricated data after performing initial data-analyses

Sample exclusion or extension to obtain the desired outcome

Searching for a statistical test or statistical package that gives the most significant result

Hiding important information from the data with wrong statistical test

Deleting or removing some part of the data set to get a nice curve or a nice result

Intentional destruction of archaeological or otherwise historical artefacts

Adjust covariates to get positive results

Selectively including/excluding data without disclosure

Adjust the study's experimental power until the desired effect can be reported

Intentionally misinterpreting results from a statistical analysis

22. ~~Perform data analyses not stated in the study protocol without disclosure~~

Perform post-hoc analyses without disclosure

23. Report an incorrect downwardly rounded p-value

24. Not report all study protocol-stipulated results

Failing to disclose all measures collected in a study

Selectively choosing primary outcome study

Present secondary outcomes as primary if initial primary outcome gives no or little result

25. Not publish a valid 'negative' study

Only considering to publish a study when the results are negative

Not publishing results

26. Report an unexpected finding as having been hypothesized from the start

27. Conceal results that contradict earlier findings or convictions

28. Not report clearly relevant details of study methods

Changing the parameters in the method so that other researchers cannot reproduce your results

Not sharing all details relevant for reproducing experimental results

Not being transparent in reporting

29. ~~Not report replication problem~~

Not reporting of failure to replicate

30. Selectively cite to enhance own findings or convictions

Intentionally use a selective bibliography

31. Selectively cite to please editors, reviewers or colleagues

Referees that impose citations to their own work

32. Selectively cite or cite own work to improve citation metrics

33. Let own convictions influence the conclusions substantially

Treating an opinion as more valuable than evidence from adequate research.

Lack of doubt or being too convinced about own research to soundly judge its merit

Flexible interpretation of the literature so it fits your hypothesis

34. Insufficiently report study flaws and limitations

35. Spread study results over more papers than needed

Salami slicing: To split one research question/data set into multiple articles to publish more with the same work

Cutting up data and spreading it over as many publications as possible

36. Duplicate publication without disclosure

Submitting/publishing the same material (text) in different journals chapters

37. Re-use parts of own publications without referencing

Self-plagiarism

38. Re-use of previously published data without disclosure

“Double publishing” of similar research more than once

39. Modify the results or conclusions of a study due to pressure of a sponsor **or other stakeholder**

40. Failure to disclose a sponsor of the study

41. Failure to disclose a relevant financial or intellectual conflict of interest

42. Handle existing conflicts of interest inadequately

43. Communicate results to the general public before a peer reviewed publication is available

44. Deliberately communicate findings inaccurately in the media or during presentations

Intentionally misrepresenting the scientific literature to the public

45. Make no clear distinction between personal views and professional comments

46. *Investigating the null hypothesis less thoroughly than the alternative hypothesis*

47. Exaggerate experimental results

Interpreting results inadequately and present them as spectacular effects with little critical evaluation

Oversell research as more than it is

48. No critical appraisal of negative aspects of novel methodology

49. Writing a paper too fast without proper and rigorous analysis

Claiming that something is new or offers a new perspective without proper reading

50. ***Change a manuscript due to pressure from editors or reviewers in a grossly***

incorrect way in order to get it published

51. *Overutilization of databases*

52. *Exclusion of outliers without disclosure*

Deleting or removing some part of the data set to get a nice curve or a nice result

Collaboration

53. Take no full responsibility for the integrity of the research project and its reports

54. Refuse to share data with bona fide colleagues

Withholding data for reproduction of published work

Treating data as if you “owned” it and other researchers have no rights to use that data

55. Turn a blind eye to putative breaches of research integrity by others

56. Refuse to respond to an allegation of a breach of research integrity

57. Use unpublished ideas or phrases of others without their permission

Referees that steal ideas from proposals or manuscripts they need to evaluate

As a reviewer rejecting a paper and then stealing the ideas from that paper

*Taking PhD students' or students' work for publication without
permission/acknowledgement*

58. Use published ideas or phrases of others without referencing

Steal data or copy findings without consent

Plagiarism

59. Unfairly review papers, grant applications or colleagues applying for promotion

Nepotism in grant funding, publication review, promotion, etc.

*To accept papers for conferences or articles for proceedings based on names or
someone's status in the field*

Feeling you are obliged to accept a bad paper because it is someone's PhD student

Declining or withholding publication from a competitor

Abuse of referee power

60. Review one's own submitted manuscripts

61. Demand, accept or offer substantial gifts for doing a favour

62. Insufficiently **or inadequately** supervise or mentor junior coworkers

*Knowingly supervising a PhD student whilst there is not sufficient knowledge on their
topic in the research group and failing to direct them elsewhere*

Pressuring junior postdocs and PhDs in "finding" significant effects

*Pressure your PhD student to work on other research projects while PhD student is
still working on former (PhD) project*

*Not provide a safe learning environment where they are allowed to learn and make
mistakes for more junior researchers*

Supervisors writing (part of) chapters for PhD students to guarantee graduation

Intentionally wrong role modelling for junior colleagues to create a culture of fear

63. Gross unfairness to collaborators

Failing to attribute own success to team effort

Not doing your fair share in collaborations

64. Add an author who doesn't qualify for authorship

Guest authorship

First authorship or co-authorship without doing significant work

65. Demand or accept an authorship without significant contribution

66. Omit a contributor who deserves authorship

Ghost writing a paper for someone else

67. Not acknowledge contributors who do not qualify for authorship

68. Not ask permission from contributors for the wording of the acknowledgement

69. Not share reviewers' comments with all co-authors

70. Submit or resubmit a paper or grant application without consent from all authors

71. ***Misuse of research grants for unintended purposes***