Author’s response to reviews

Title: Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity

Authors:

Lex Bouter (lm.bouter@vu.nl)

Joeri Tijdink (j.tijdink@vumc.nl)

Nils Axelsen (na@ssi.dk)

Brian Martinson (brian.c.martinson@healthpartners.com)

Gerben ter Riet (g.terriet@amc.nl)

Version: 1 Date: 14 Sep 2016

Author’s response to reviews:

Amsterdam, September 14th, 2016,

Dear Dr Harriman,

Thanks for your fast and constructive comments on our manuscript "Ranking major and minor research misbehaviors: results from a survey among participants of four World Conferences on Research Integrity" (RIPR-D-16-00015). Below we respond to them point-by-point, also indicating the corresponding changes we made in our article. The page and line numbers mentioned refer to our original manuscript. We uploaded the revised version of our manuscript - both with and without the changes visible - to your manuscript management system as requested.

All authors contributed to the revision and agreed with the final revised version that was submitted.

We believe that the revised manuscript is substantially improved and hope that you will now accept it for publication in Research Integrity and Peer Review.

Kind regards,

Lex Bouter
In addition to the comments from the reviewers, please address the following Editorial comments.

1) The response rate is reported as 77% based on the number invitees who opened the web-based questionnaire and 33% based on the number of invitees who opened the invitation email. Please report the response rate based on overall participants who were sent the invitation email - 17% (227/1345).

Reply – We now report a response rate of 17% in the abstract and mention in the Results and Discussion sections the alternative ways to calculate the response rate. Table 1 now shows the corresponding calculations.

P2 - L12-13: Response was 77% among invitees that opened the web-based questionnaire and 33% among those who opened the invitational e-mail. Response was 17% of those who were sent the invitational email and 33% of those who opened it.

P6 – L22-23: The response rate was 77% among invitees that opened the web-based questionnaire, but only 33% among those who opened the invitational e-mail (Table 1). Response was 17% of those who were sent the invitational email, 20% of the valid email addresses, 33% of the emails opened and 77% of the links to the survey opened (see table 1).

P10 – L8-10: The response rate of 77% of our study was adequate when focusing on the invitees that opened the web-based survey. But we were disappointed that only 42% of those who opened the invitational email opened the survey. The response (17%) was low, although some email addresses were invalid and only about half of the emails were opened. We were disappointed that only 42% (293/693) of those who opened the invitational email also opened the survey.

Table 1. Response rates to the survey

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Response rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Invitational e-mails</td>
<td>1,345</td>
<td>17% (227/1,345)</td>
</tr>
<tr>
<td>Valid email addresses</td>
<td>1,131</td>
<td>20% (227/1,131)</td>
</tr>
<tr>
<td>Emails opened</td>
<td>693</td>
<td>33% (227/693)</td>
</tr>
<tr>
<td>Links to survey opened</td>
<td>293</td>
<td>77% (227/293)</td>
</tr>
<tr>
<td>Participants that completed survey</td>
<td>227</td>
<td></td>
</tr>
</tbody>
</table>

2) Please revise to follow the STROBE guidelines.
Reply – We applied the STROBE checklist for cross-sectional studies to the revised manuscript and noticed that our article complies with all applicable recommendations.

3) We note that you included a more detailed statement regarding ethics in your answers to the questions on submission. Please can you include this fuller statement in the declarations section of your manuscript. For your reference, I have cut and pasted the statement you gave on submission:

"Under the pertaining laws and (institutional) regulations in The Netherlands no permission from a (medical) ethics committee (Institutional Review Board) is needed. We sent an invitational email containing a link to the web-based survey to participants of one or more participants of the four past World Conferences on Research Integrity. Invitees gave their consent by following this link and by submitting their answers to the web-based questionnaire. Invitees could also opt-out with another link in the invitational email. These invitees got no reminders."

Reply – We included the fuller statement in the revised manuscript as requested.

P13 – L6 - 10: Not applicable. Under the pertaining laws and (institutional) regulations in The Netherlands no permission from a (medical) ethics committee (Institutional Review Board) is needed. We sent an invitational email containing a link to the web-based survey to participants of one or more participants of the four past World Conferences on Research Integrity. Invitees gave their consent by following this link and by submitting their answers to the web-based questionnaire. Invitees could also opt-out with another link in the invitational email. These invitees received no reminders.

Reviewer #1:

1) Why wasn't this survey reviewed by any institutional review board? My institution would have wanted to review it, even if the final conclusion was that it was exempt.

Reply – See our reply to editorial comment 3. We now write:

P13 – L6 - 10: Not applicable. Under the pertaining laws and (institutional) regulations in The Netherlands no permission from a (medical) ethics committee (Institutional Review Board) is needed. We sent an invitational email containing a link to the web-based survey to participants of one or more participants of the four past World Conferences on Research Integrity. Invitees gave their consent by following this link and by submitting their answers to the web-based questionnaire. Invitees could also opt-out with another link in the invitational email. These invitees received no reminders.
2) Page 5, line 6, the number of valid email addresses does not agree with that in Table 1.

Reply – Thanks for pointing this out as it’s indeed confusing. We meant to explain that 8 email addresses were clearly incorrect and were not used. In Table 1, valid email addresses mean that the message did not bounce and was successfully delivered in the mailbox at issue.

P5 – L11: 1345 after omitting invalid clearly incorrect email addresses.

3) Page 5, lines 10-11, how were the 60 items assigned to topic domains? These domains should be described in more detail, as well.

Reply – We provided some more detail in Additional file 3 of the revised manuscript.

Additional files page 5: The authors interactively reached consensus on the allocation of the items to one of four topic domains: study design (items that concern the phase before the start of data collection), data collection (items that concern the phase of data collection), reporting (items that concern data-analysis and reporting the results of the study), and collaboration (items that concern obligations towards colleagues and science as a whole).

4) Page 5, lines 7-11, the sampling description is very confusing. On the one hand it sounds like each participant was given an individual random selection of 20 items, but then the 20 items were selected by stratified random sampling. Yet, there were 3 groups of participants for some reason. This needs to be described more coherently.

Reply – We tried to describe this more clearly.

P 5 – L13-16: …..were randomly allocated to 3 groups [18]. We randomly allocated 20 out of 60 items to each group after stratification on the topic domains (study design, data collection, reporting and collaboration) (see Additional file 3). Each participant received an email message (see Additional file 1) with a link to a questionnaire (see Additional file 2) containing a the random selection of 20 out of a total of 60 items of research misbehavior for that group. (see Additional file 3).We randomly allocated the items to the 3 groups after stratification on the topic domains of: study design, data collection, reporting and collaboration (see Additional file 3). For each participant we randomized the order of presentation of the 20 items.

5) Page 6, lines 4-7, the various ways in which the items were ranked should be described.

Reply – We agree. We have extended the text where we describe the ways in which the items were ranked.
P6 – L10-15: Data of the 60 items on frequency, impact on truth and trust, and preventability were exported from SurveyMonkey in SPSS format. Data were then read into STATA 13.1 using StatTransfer (version 12-64). Simple descriptive analyses were performed using STATA 13.1. The ranking was based on the mean scores on frequency, impact on truth and impact on trust of the 60 items. The item means of the products of the scores for frequency and for impact on truth, of frequency and impact on trust, and of frequency and preventability were also calculated and the items were then ranked accordingly using STATA 13.1.

6) Page 6, line 12 states data are not shown for reasons to decline participation. Can these data be summarized?

Reply – We have summarized the results of the non-response survey. We now add:

P7 – L2 – 5: 33 invitees (5% of those who opened the invitational e-mail) disclosed their reasons for declining participation and provided basic background information. 39% of them argued that they were not involved in the area of research on research misbehaviors, while 24% reported having no professional link to the area of research integrity.

7) Page 6, lines 18-19, why is it relevant to add a comment on the percentage of South American participants?

Reply – We wanted to provide the likely explanation for the high proportion of respondents from South America. But prompted by your remark we decided to remove the sentence from the manuscript.

P7 – L11-12: Twenty percent came from South America, but that is most likely because the most recent World Conference on Research Integrity in 2015 was held in Rio de Janeiro, Brazil.

8) Page 9, lines 7 - 8, it is likely that a large percentage of respondents selected 'Other' because the list of options over-represented academic positions and under-represented positions in other fields.

Reply – We agree that this is indeed a likely explanation and added a remark on it.

P9 – L19-23: It is interesting to note that more than a quarter (27%) of our respondents reported their job title as ‘Other’, which might be due to an overrepresentation of academic job titles.
among the answer options. Therefore, we know little about the ‘credentials’ of this group pertinent to the survey questions.

9) Table 3, for the frequency section, the 5th item is listed as in domain (C), but Supplemental file 3 puts this item in domain (R). There is a similar problem for the 5th item in the trust section of Table 3. I did not check all of the items in Table 3, but the authors should double-check their tables.

Reply – Thanks for pointing out these mistakes. We corrected them and double-checked all tables. We found three more errors in Table 3 that concerned the same item.

P15-18 - Table 3:

Selectively cite to please editors, reviewers or colleagues (CR)

Unfairly review papers, grant applications or colleagues applying for promotion (RC)

Use published ideas or phrases of others without referencing (RC)

Use published ideas or phrases of others without referencing (RC)

Use published ideas or phrases of others without referencing (RC)

10) Consider putting Table 3 in the Supplementary Material section and replacing it graphically in the main body of the paper with bar charts including error bars representing the 95% CIs, and, perhaps coloring or shading to indicate domains.

Reply – We have considered using bar charts. We decided to stick to the tables as within the bar charts the text labels become quite long which hamper the readability and interpretability of bar charts.

11) Supplemental file 1 - the letter of invitation contains wording that may bias participants. For example, in the second paragraph, the authors suggest that questionable research practices may do more harm than intentional deceit. This may influence participants to give answers that support the authors' claim.

Reply – In the invitational email we tried to explain as clearly as possible what the aim of our project was and what we wanted the participants to do. We did not want to influence the participants by the sentence ‘But perhaps ‘sloppy science’ due to questionable research practices does even more harm on the aggregate than intentional deceit.’ This was merely meant as a
description of our study question. But we cannot exclude that the phrasing of our invitation influenced the responses to the survey. We added a remark in that sense to the discussion section.

P10 – L22 - P11 L 1 – 2: Furthermore, although we tried to formulate the 60 items and our invitational email and the survey questions and answer options as unambiguously as possible, we do not know if the respondents, with their wide range of disciplinary backgrounds, had a common understanding of the issues raised and were not unduly influenced by the way we formulated the invitational email and the survey questions.

12) Supplemental file 2 - question 8, in a sense, assumes that all participants have a PhD. What if the participant had a degree other than a PhD, such as an MD, DVM, DrPH, etc. or a masters or bachelors degree?

Reply – This question only concerns PhD degrees. Respondents with no degrees or other degrees had to tick the box ‘not applicable’.

13) Supplemental file 3 - were the additional comments within brackets, [ … ], made available to the participants? If so, they may influence the responses unfairly, compared to items without comments.

Reply The additional comments in brackets were made available to the survey participants. They were included to be explanatory to make sure the participant understood the exact meaning of the item. We thought the items without comments were self-explanatory.

14) Supplemental file 4 - provide sample sizes.

Reply - We have included the number of respondents per item in Additional file 5. The number of respondents per item ranges from 46 -61.

15) Supplemental file 5 - the table format can be cleaned up. Some of the confidence intervals have parentheses around them and most do not. The column labels for modes of prevention strategy can leave out the word, Prevention, and (%), as this information is in the top header of these columns.

Reply - We have followed your advice and have cleaned up the table format. See Additional file 5.
16) Supplemental file 6 - provide sample sizes for the disciplinary fields.

Reply – We have included the number of respondents (range) per disciplinary field in the subheadings of supplemental file 6.

Reviewer #2:

The topic is interesting. There is a definite need to go beyond falsification, fabrication and plagiarism when talking about research misconduct

1) Overall, the manuscript would benefit from being reported according to the STROBE checklist for cross-sectional studies.

Reply – We have applied the STROBE checklist to our revised manuscript. See also our reply to editorial comment 2.

2) General concerns regarding this study, some of which are well reported as limitations of the study in the discussion part, include the fact that this reflects personal opinions of a convenience sample of participants with a low response rate (227 out of 1131 invited) and over-representation of Professors who may have a different (biased) view on the prevalence of misconduct occurring during the data extraction phase of a clinical trial (and could explain the conclusions of the study regarding the importance of lack of mentorship of junior researchers, etc…)

Reply – In response to this comment we added two sentences to the discussion section.

P10 – L8-10: The response (17%) was low, although some email addresses were invalid and only about half of the emails was opened. We were disappointed that only 42% (293/693).of those who opened the invitational email also opened the survey.

P11 – L 4 – 5: We cannot exclude that these characteristics influenced the rankings, for instance when the many professors and the few PhD students in our sample would have different opinions on authorship and supervision.

Specific comments:

3) Abstract:
Method, line 7: I suggest to the reader's attention to the fact that the attendees were asked to score these items according to « their idea of the prevalence, preventability etc… » since obviously, they have no knowledge of the real prevalence.

Reply – We agree and modified the text of the abstract accordingly.

P2 – L6-7: They were asked to score 60 research misbehaviors according to their views on and perceptions of the frequency of occurrence,…...

4) Results: I suggest to report that 227 participants responded to the questionnaire, (20% of invited, 33% of those who opened the email)

Reply - See our reply to editorial comment 1. We now write:

P6 – L22-23: Response was 17% of those who were sent the invitational email, 20% of the valid email addresses, 33% of the emails opened and 77% of the links to the survey opened (see table 1).

5) Conclusion: the study says nothing on the potential impact of intervention to promote research integrity (last sentence of conclusion seems to be over-interpretation of results).

Reply – We agree and removed the last sentence of the abstract.

P2 – L18-20: Many scientists are thought to engage in sloppy science. Interventions – educational or otherwise – promoting a research culture that actively discourages cutting corners, seem to be a way forward.

We have added the sentence:

P2 – L21-23: Our respondents were much more concerned over sloppy science than about scientific fraud (FFP). In the fostering of responsible conduct of research we recommend to develop interventions that actively discourage the high ranking misbehaviors from our study.

We changed the conclusion section of our manuscript accordingly.

6) Introduction:

The very first sentence seems to contain a contradiction: « Responsible conduct of research is the norm », « breaches of research integrity occur alarmingly commonly ». 
Misconduct, misbehaviour, questionable practice are all ill-defined terms commonly read in the papers reporting on research integrity. If a questionable research practice « often (...) threaten the relevance, validity, trustworthiness or efficiency of the study (...) and may be committed intentionally (...) » why should we differ it from a misconduct?

The real question is why do scientist not agree that the behaviour is undesirable and should be avoided?

Reply – We changed the first sentence to avoid this contradiction. We agree that the key concepts are ill-defined in the research integrity literature, but do not believe that our introduction section is the right place to solve this. Our core concept is research misbehavior and we operationalized that in the form of the 60 items. But we felt we also needed to connect to some of the other terms that dominate the literature (research misconduct (FFP), questionable research practices, sloppy science).

P3 – L7: While responsible conduct of research is expected of scientists the norm [1-3], breaches of research integrity occur ……

7) Page 4, line 1 : Authors state that the « total harm » caused by a specific research misbehaviour depends on the frequency of its occurrence and the impact when it occurs. Again, the impact when it occurs may be very difficult to quantify, if you think of the Wakefield example regarding the relationship between vaccination and autism.

Reply – We agree. The impact is difficult to assess empirically and may vary substantially between instances. We added a sentence on this.

P4 – L6-7: Furthermore, the impact is difficult to assess empirically and may vary between instances.

8) Method

The method section should follow the SROBE recommendation for reporting of cross-sectional studies. Using all the subheadings should help to clarify the lacks in information provided. (study design : cross-sectional ; Setting ; Participants ; etc..)

Reply - See our reply to editorial comment 2. We have applied the STROBE checklist to our revised manuscript. We decided against using so many subheadings, but made sure that all relevant information is provided.
9) Page 4, line 16: "our first aim (...) list of misbehaviour, acknowledged by experts" how are the experts defined? Is attendance to a conference enough to be called an expert?

Reply - See our reply to comment 11 of reviewer 2 below. We now explain in the Additional file 3 that the experts involved were a) the authors, 2) the members of a committee that drafted a national research program on fostering responsible research practices in the Netherlands, and 3) keynote speakers and session chairs of the 4th WCRI.

10) Page 4, line 17: I would add the term "perceived" "(...) to provide expert-based ranking of these specific research misbehaviour according to PERCEIVED frequency of occurrence..."

Reply – See our reply to comment 3 of reviewer 2. We made the same textual change here.

P4 – L22: Our second study aim is to provide expert-based rankings of these specific research misbehaviors according to their views on and perceptions of the frequency of occurrence.....

11) Page 4, line 22: how many candidate items did you start with? Which survey questions? Which relevant literature (references?) this paragraph is a bit too vague to be informative.

Reply – We added a more detailed explanation of the way we arrived at our 60 items list to Additional file 3.

Additional files page 5: This list of items was derived at in five steps. Firstly, a comprehensive list of more than 100 items and sub-items was compiled from the literature, codes of conduct, guidelines, and our own experiences. This first list covered all the major and minor deviations from responsible conduct of research we could think of. Secondly, we shortened the list to 60 items by eliminating redundancies, merging sub-items, rephrasing some of the items, and reformulating a number of items to make them relevant for all empirical research. Thirdly, this list was tried out by a group of 15 experienced researchers drafting a national research program on fostering responsible research practices in the Netherlands. In an interactive workshop the items were ranked and comments on the formulation and the relevance of the items plus suggestions for items that also should be included were made. This resulted in some changes in the item list. Fourthly, we invited a selected group of 60 keynote speakers and session chairs 6 months before the 4th World Conference on Research Integrity (4th WCRI), Rio de Janeiro, 2015. In a web-based survey they were asked to estimate frequency of occurrence and impact on truth and trust of all items, to suggest lacking items, and on the formulation of the items. Fifthly, the 34 respondents were invited (26 attended) for a workshop at the 4th WCRI which resulted in the list of items used in the project reported on in this publication. Also the study design and the formulations of the survey questions and answer options were based on the workshop.
12) Page 5: additional file 1: What happened to Daniele Fanelli, who signed the invitation later? Did his contribution stop at this stage? (this is just by curiosity…)

Reply – When reading the penultimate draft of our manuscript Daniele Fanelli came to the conclusion that did not contribute enough to qualify as co-author. With his consent we refer to his contribution in the acknowledgement section.

13) How were the « colleagues » selected? were all speakers and session chairs of the conference invited? Or how was the selection performed? How many attended the workshop? On what basis were the 60 items selected? Out of how many?

Reply - See our reply to comment 11 of reviewer 2.

14) How did you get the list of the participants of the 4 last congress? More detailed information would be interesting.

Reply – We used the combined list of email addresses of the participants who attended one or more World Conferences on Research Integrity that was constructed to announce the 5th WCRI in Amsterdam, May 2017.

15) Page 5: the sentence regarding "the emphasis that questions pertain to views of the respondents, and their subjective opinion of unknown basis" should come either earlier in the introduction/objective.

Reply – See our replies to comments 3 and 10 of reviewer 2. In response to these comments we now emphasize the subjective nature (views, perceptions) two times earlier in the revised manuscript.

16) Page 6, table 2:

A brief description of the reasons provided by the 33 participants not willing to answer the questionnaire could be interesting.

Over-representation of professors could be highlighted in the text.

Reply - See our replies to comments 6 of reviewer 1 and 2 of reviewer 2. We now write:
33 invitees (5% of those who opened the invitational e-mail) disclosed their reasons for declining participation and provided basis background information. 39% of them argued that they were not involved in the area of research on research misbehaviors as main reason for non-participation, while 24% reported no professional link to the area of research integrity.

We cannot exclude that these characteristics influenced the rankings, for instance when the many professors and the few PhD students in our sample would have different opinions on authorship.

Overall, the interpretation of the 95% CI in the tables is unclear. Are we drawing conclusion on "the whole community of "experts" attending a given conference and that agree to participate in an online survey"? In which case, I'm not so sure that this "confidence" reflects much.

Reply – We agree that the interpretation of 95% CI in a convenience sample with a low response rate is a bit problematic. But it’s not unusual to use 95% CI as an indicator of precision in situations other than inferences from a random sample to the source population.

Finally, while revising our manuscript we also made some manor changes to sentences that were not commented on by the editor and the reviewers.