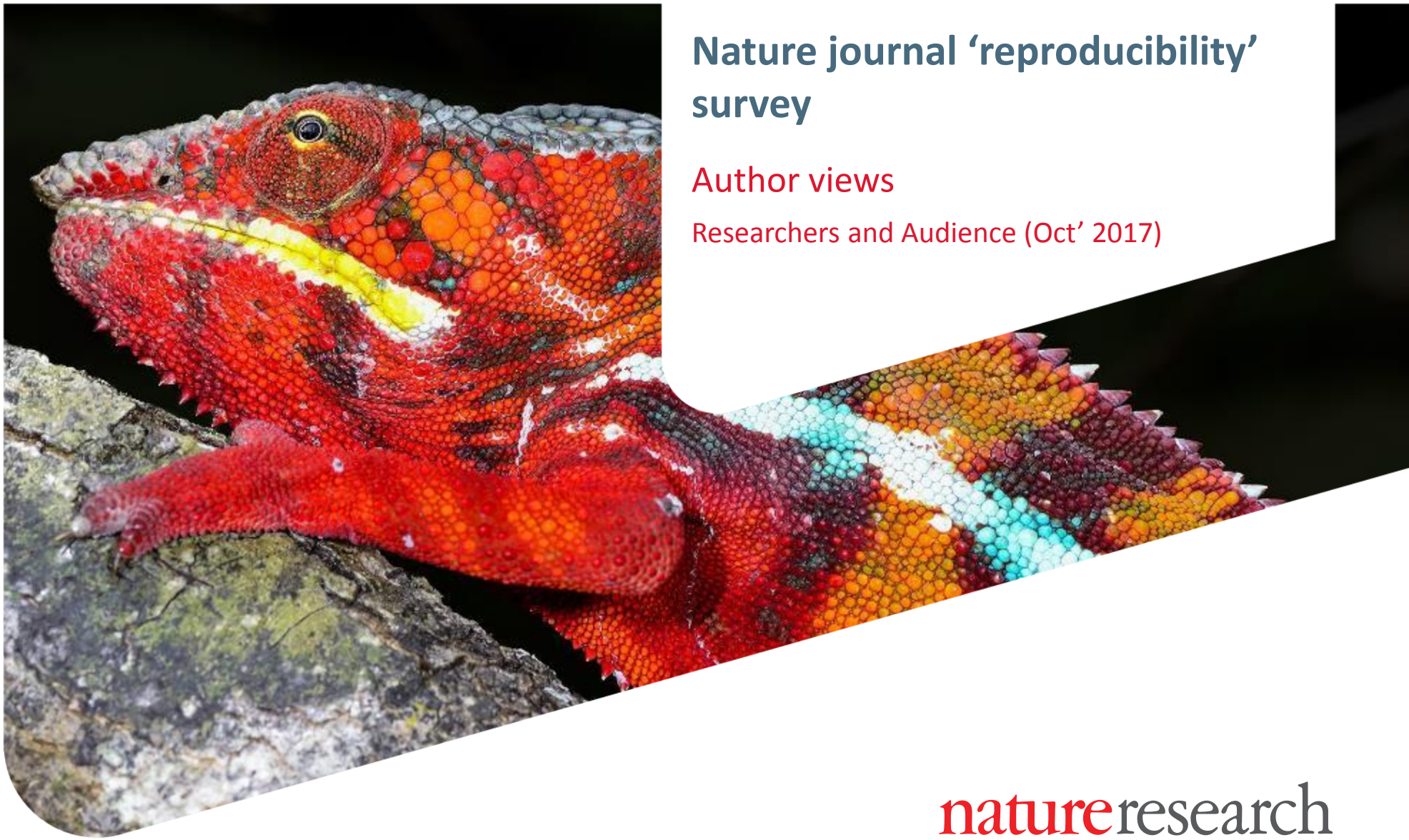
SPRINGER NATURE



How chameleons change colour

Nature journal 'reproducibility' survey

Author views

Researchers and Audience (Oct' 2017)

nature research

Key Findings

- 86% of respondents felt that there was a 'crisis of reproducibility' within their field of research.
- In terms of where they had heard about the crisis, it appears that a large proportion are talking about it in laboratories amongst their colleagues with 58% indicating they had heard of it through this mode.
- Selective reporting of results was selected by the greatest proportion of respondents (66%) as the factor that contributed to irreproducible research. As such, 58% of respondents placed researchers as the most influential stakeholder in improving the reproducibility of research.
- Action taken by publishers were deemed to be having the greatest effect on reproducibility in comparison to initiatives started by research institutions and funding agencies. However, three quarters of panellists also felt publishers should do more to encourage reproducibility.
- The Nature journal reporting checklist is considered by 50% of first time submitters only after the first draft of their manuscript is submitted.
- Encouragingly, 78% of respondents continue to implement the checklist to some degree with their continued research.
- While 43% of respondents felt that the checklist had led to better reporting of statistics, there were another 43% who were unsure as to whether the checklist had improved the quality of research in their field. The latter group did not feel the checklist addressed the real issues of reproducibility.

- Considerations – The results from this survey provide a useful snapshot of the opinion of Nature journal authors. However, it needs to be considered that the cohort of respondents are very westernised with only 14% coming from outside of Europe and the US.

Introduction and Methods

Introduction

Nature journals have been openly discussing and publishing news and comment with regards to the area of 'reproducibility in science' since 2010. From then, multiple initiatives have been implemented including roll out of a reporting checklist (2013) for all submitted manuscripts; and setting up a Reproducibility Steering Group to deliver a roadmap to better practice. In addition to the programmes, a number of projects have been undertaken to investigate the views and beliefs of scientists regarding the status of reproducibility most notably a survey published in *Nature* (2016).

It was desired that as part of the business' continued efforts towards good practice, a survey be conducted to investigate Nature journal authors' experiences with the reproducibility-related efforts. Specifically, the survey aimed to understand:

- Current beliefs as to a 'reproducibility crisis'
- Perceived factors that contribute to a failure to reproduce results
- Perceptions of efforts made by funders, institutions and publishers alike
- Nature authors' views of the reporting checklist implemented in 2013

Method

In order to determine the perceptions of authors to the efforts of Nature journals to ensure research published is reproducible, an online survey was designed and hosted on our registered software, Qualtrics.

The survey was live from 4th September 2017 to 1st October 2017.

The survey was sent to 5,375 who had published with a Nature journal between July 2016 and March 2017. The invitation received 612 click throughs of which 480 authors completed the survey. This equates to around a 9% completion rate.

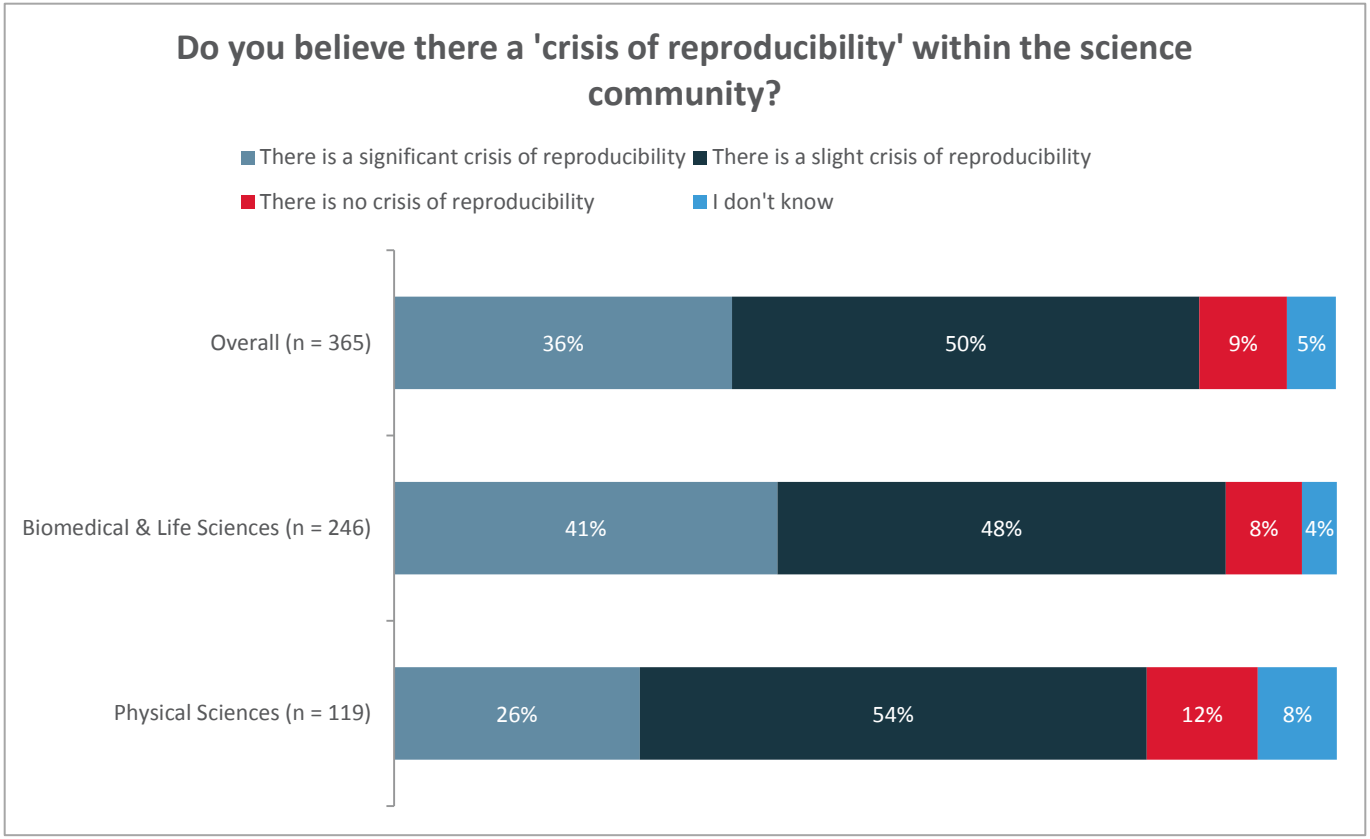
Unless otherwise stated, 480 is the base size for graphs throughout the slides.

Perceptions regarding the reproducibility of research

1.0

Crisis of Reproducibility

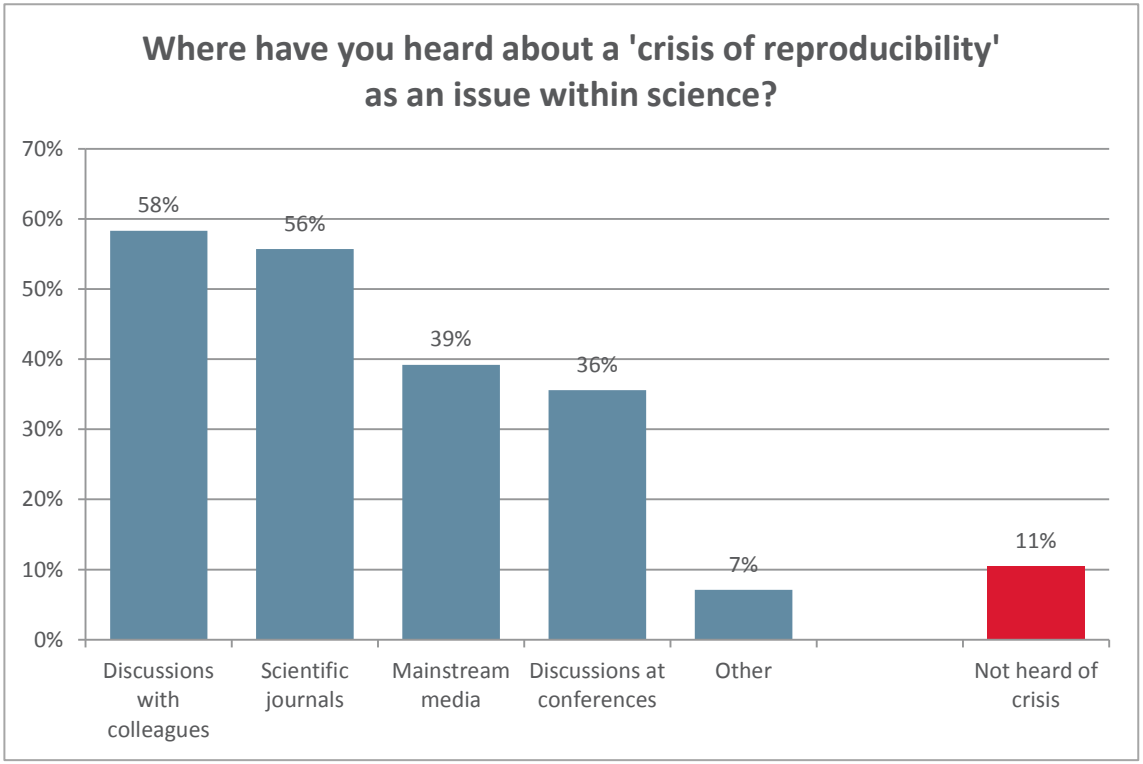
- A data point from the Nature 2016 survey indicated that 70% of scientists surveyed had tried and failed to reproduce another scientist's experiments. Unsurprisingly therefore, when asked about whether they there was a 'crisis of reproducibility' within the sciences in the current survey, 86% of respondents indicated that they felt that there was some degree of crisis.
- Interestingly, a significantly greater proportion of those in the Life Sciences (41%) felt there a significant crisis in reproducibility in comparison to those in Physical Sciences (26%) potentially reflecting the scale of the issue in each field.



Nature journal 'Reproducibility' efforts (Oct' 2017)

Where is reproducibility being discussed?

- Importantly, the largest proportion of respondents (58%) when asked about where they heard about a crisis of reproducibility, indicated that they themselves are discussing the issues of reproducibility they are experiencing in their day to day work with colleagues. Similarly, another 36% of respondents indicated that these issues are also being discussed at conferences.
- It is worth noting that 11% of respondents had not heard of a crisis of reproducibility and this roughly corresponds to the proportion of respondents who did not feel there was issue within their community (slide 3).
- Of those who selected 'Other' social media and blogs made up 45% of responses while first hand experience accounted for a further 24%.



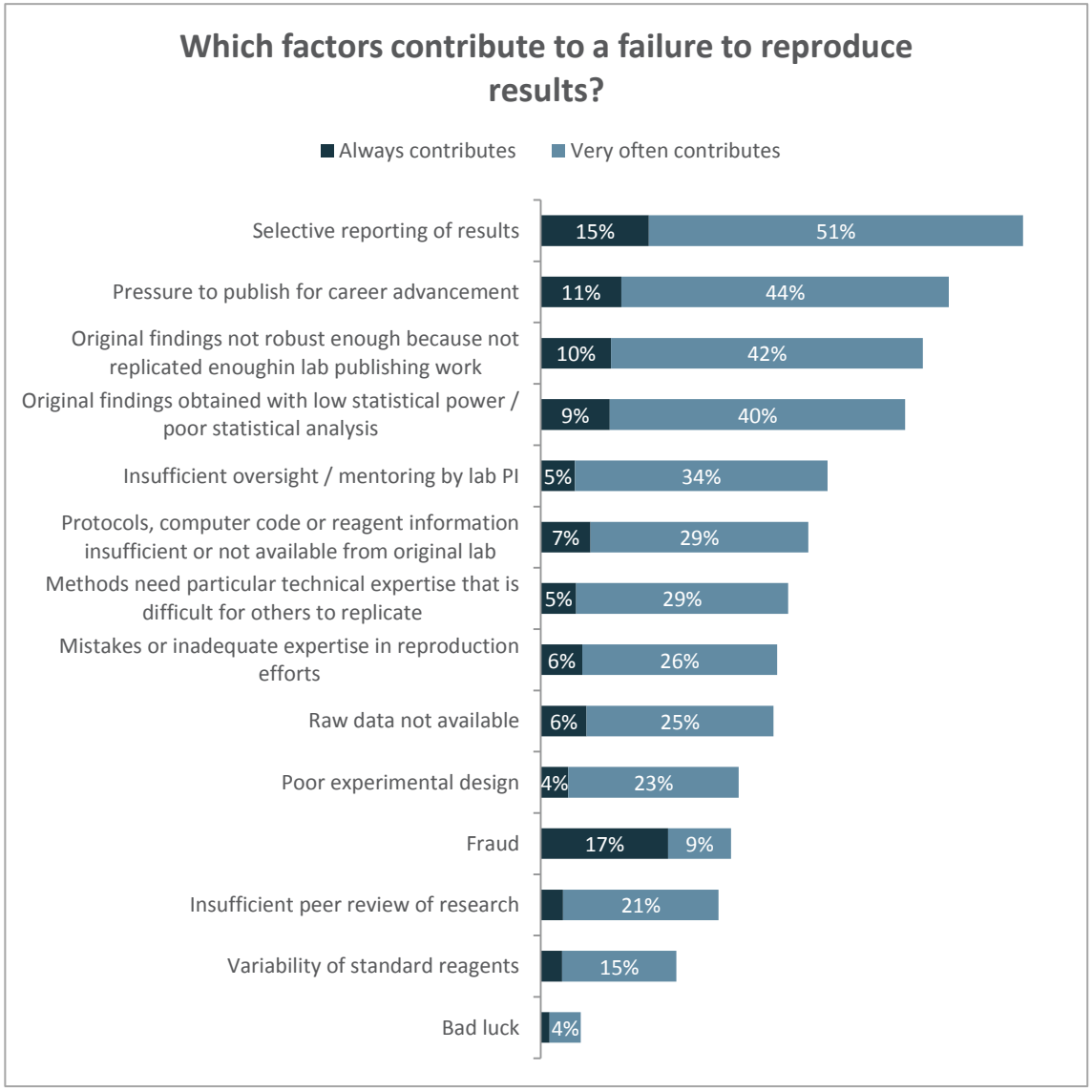
Nature journal 'Reproducibility' efforts (Oct' 2017)

Factors and influencers of reproducibility

2.0

Factors influencing the failure to reproduce others' experiments

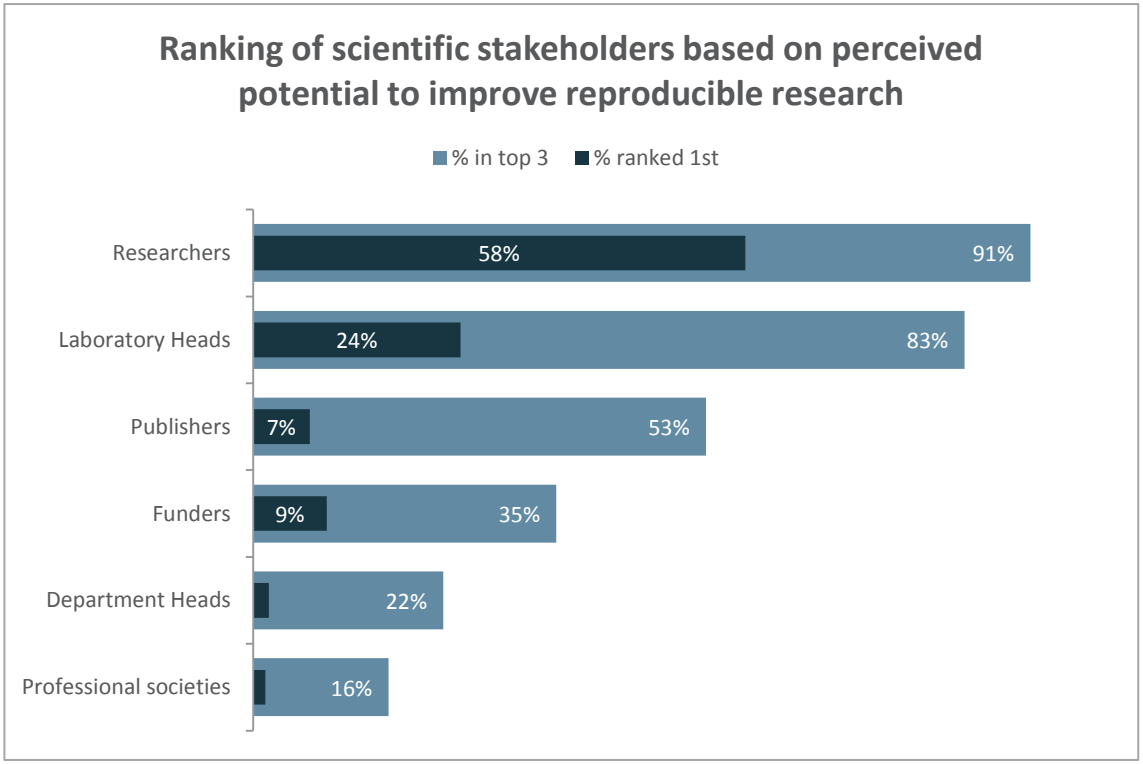
- When asked what factors manifest in irreproducible results in published work, two thirds of respondents indicated that they felt selective reporting of results was a key reason for failure.
- Including selective reporting, 3 of the top 4 selected factors relate to the quality of the original work conducted by scientists themselves leading to issues with reproducibility.
- Pressure to publish for career advancement was selected by 55% of respondents as a factor that contributes to the failure to reproduce results. As a caveat to this point, it is not clear at which stage – first publication or retesting of results – the pressure is causing bad practice.
- Interestingly, roughly a quarter a respondents felt fraud contributed to reproducibility issues and of these 65% felt it always contributed. This second proportion is significantly larger than that for any other factor.
- Some respondents felt that the survey had missed the true problems of reproducibility (see slide 23)



Nature journal 'Reproducibility' efforts (Oct' 2017)

Key stakeholders for change

- The graph below illustrates the results from a question asking respondents to rank stakeholders in science in order of those who have the greatest capacity to change and improve the issues relating to reproducibility of published work.
- In accordance with the previous slides, respondents felt that those in the labs had the greatest potential to improve the reproducibility of published work - overall 82% ranked researchers or laboratory heads as the top instigator of reproducible research with 91% of respondents ranking researchers alone in their top 3 stakeholders who had the greatest potential to improve reproducibility of published results.
- Interestingly, more respondents ranked publishers (53%) in their top three stakeholders than funders (35%) for improving the quality of reproducible research.



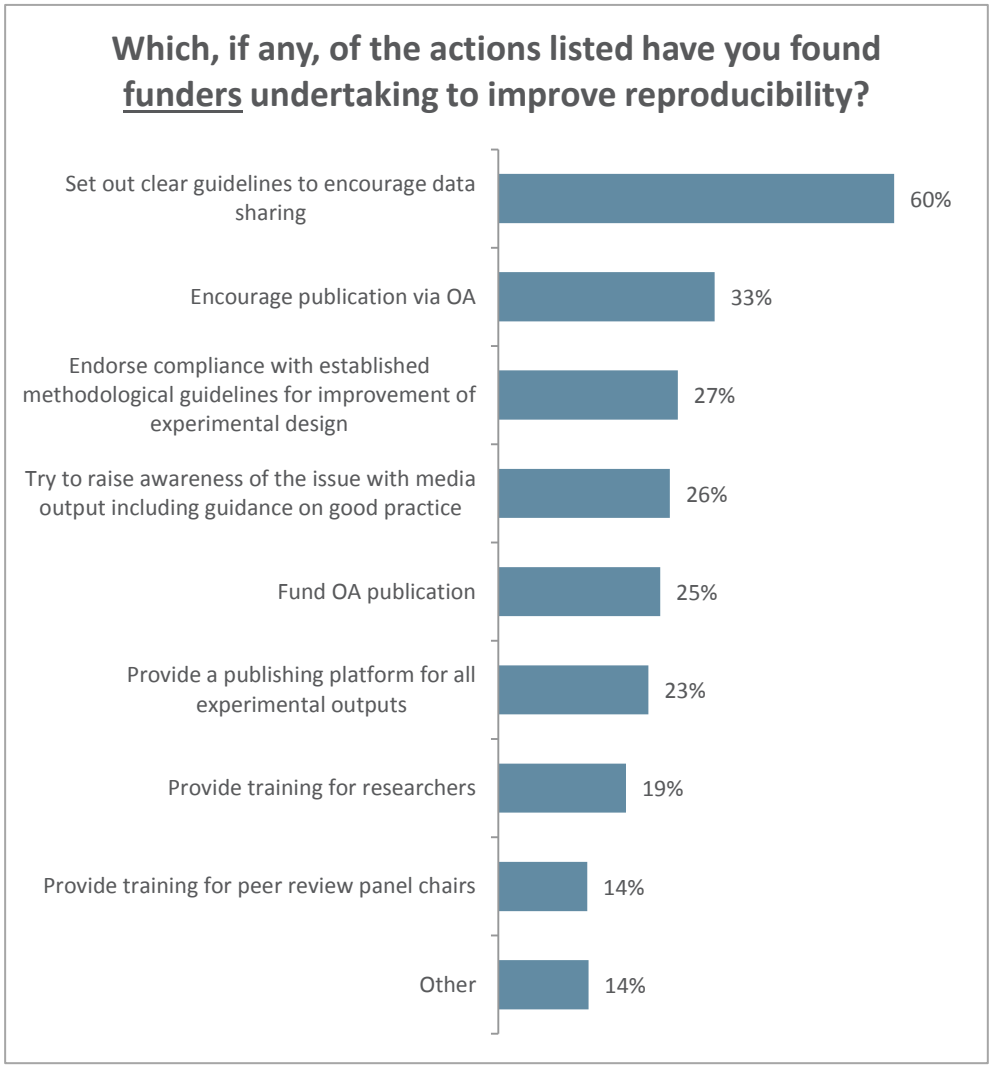
Nature journal 'Reproducibility' efforts (Oct' 2017)

Known reproducibility initiatives and effectiveness

3.0

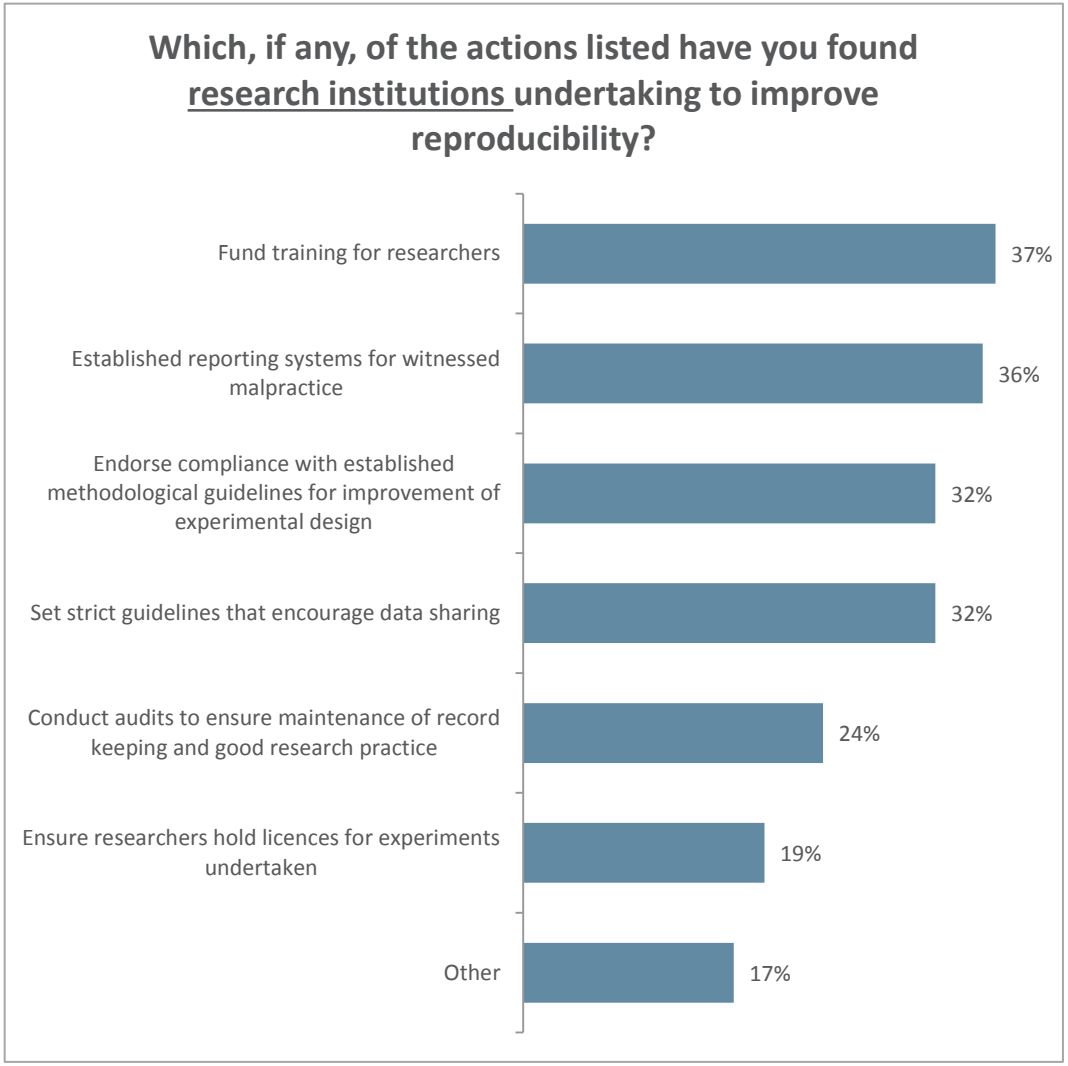
Funders actions to address reproducibility

- The standalone action that the largest proportion of respondents are familiar with funders instigating in order to improve the reproducibility of research is setting out clear guidelines to encourage data sharing (60%).
- The remaining actions listed were selected significantly less frequently as known actions funders are undertaking to try and address the issue of reproducibility, perhaps indicating that very few are aware that they undertake initiatives in this area.
- It is of note that a few respondents did leave comments indicating that they were not aware of what OA (Open Access) meant perhaps reducing its likelihood of selection. It could also indicate that researchers are not aware of particular publishing models or alternatively, they are not aware of open access in particular.



Nature journal 'Reproducibility' efforts (Oct' 2017)

Institutional actions used to address reproducibility

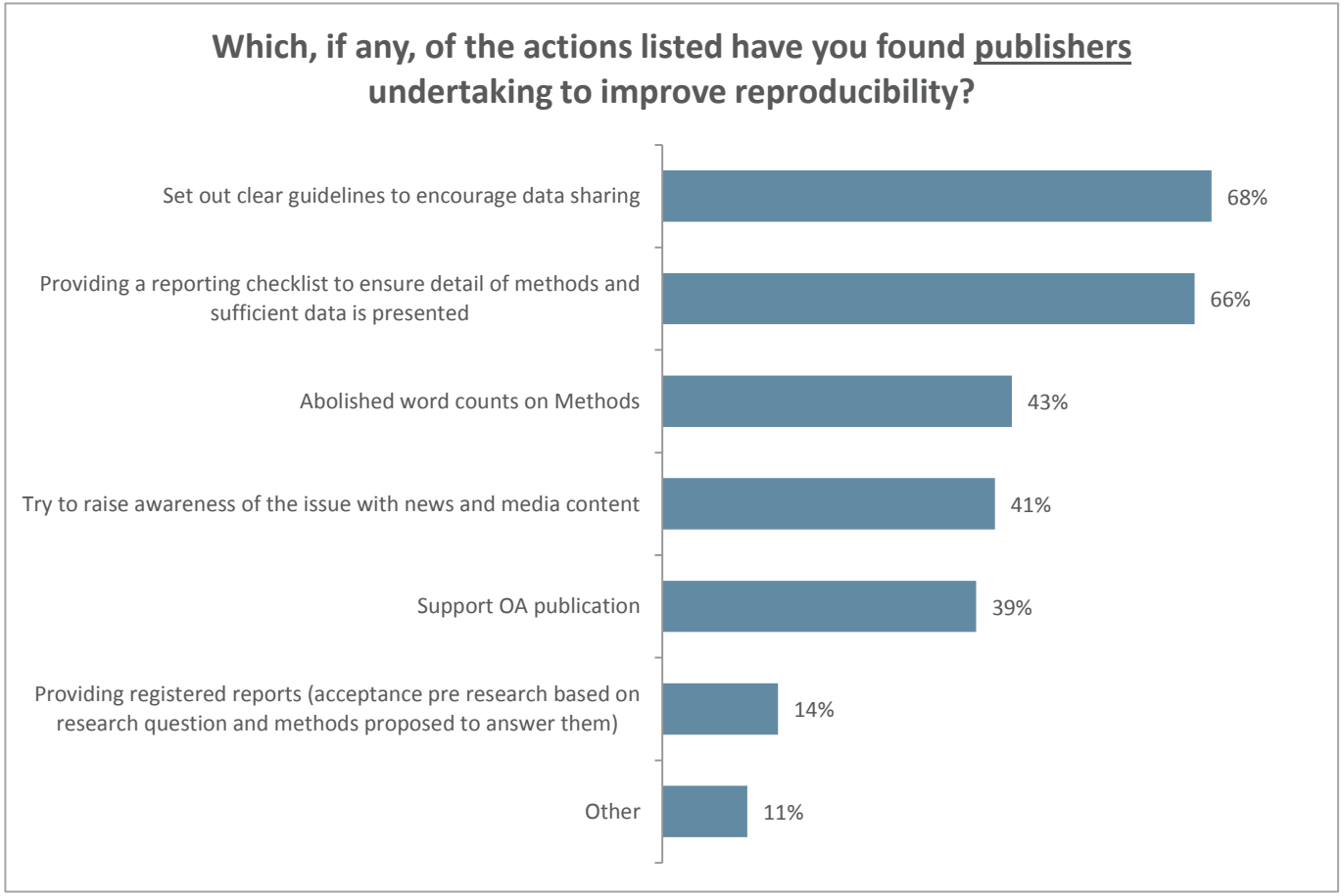


- Of the listed actions taken by institutions to address the issues of reproducibility, providing funding for training and establishing a reporting system for witnessed malpractice were both selected by over a third of respondents.
- However, the relatively low selection of any listed option indicates either that awareness of institutional initiatives for reproducibility is low or that there are few universities from which the panel resides that are trying to address this issues. Open text comments would support the later, with 64% of those who selected 'Other' referencing that they had not witnessed any actions taken by institutions to address the crisis of reproducibility.

Nature journal 'Reproducibility' efforts (Oct' 2017)

Publisher actions to address reproducibility

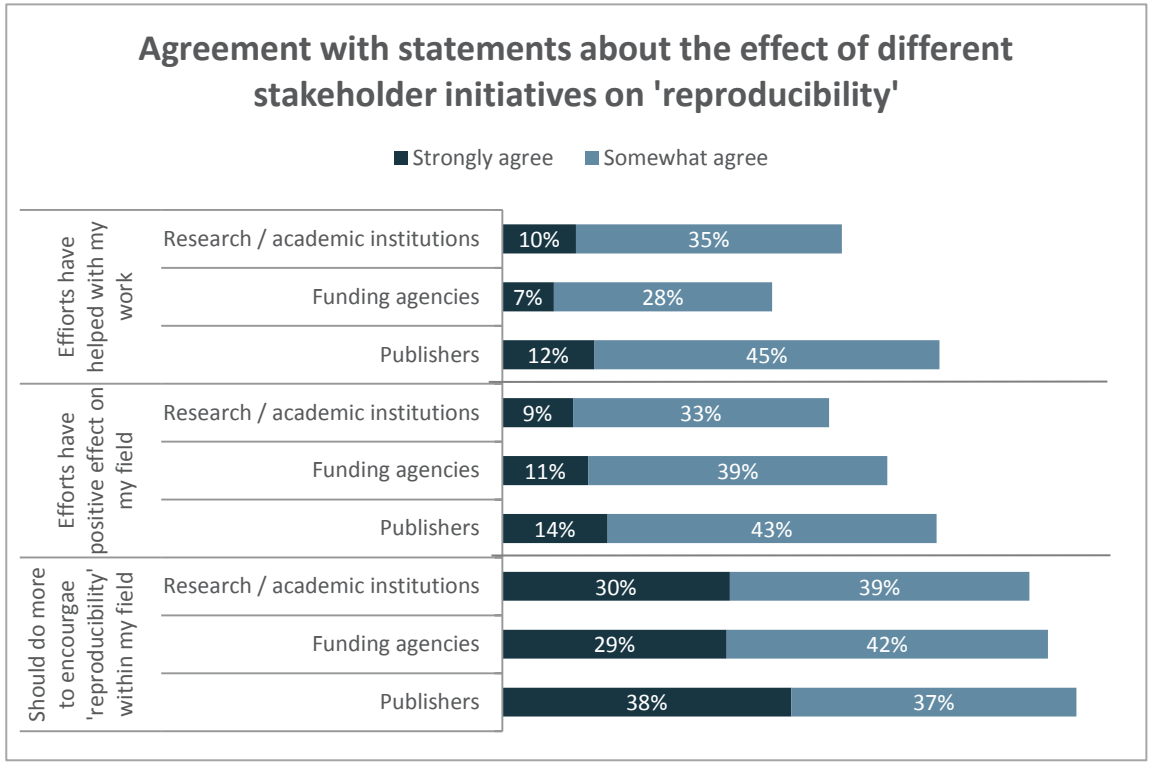
- In terms of the actions taken by publishers to try and improve the state of reproducibility of research, those best recognised by respondents were the guidelines set out to encourage data sharing (68%) and use of reporting checklists to ensure detailed methods are presented (66%).



Nature journal 'Reproducibility' efforts (Oct' 2017)

Overall perceived effectiveness of reproducibility initiatives

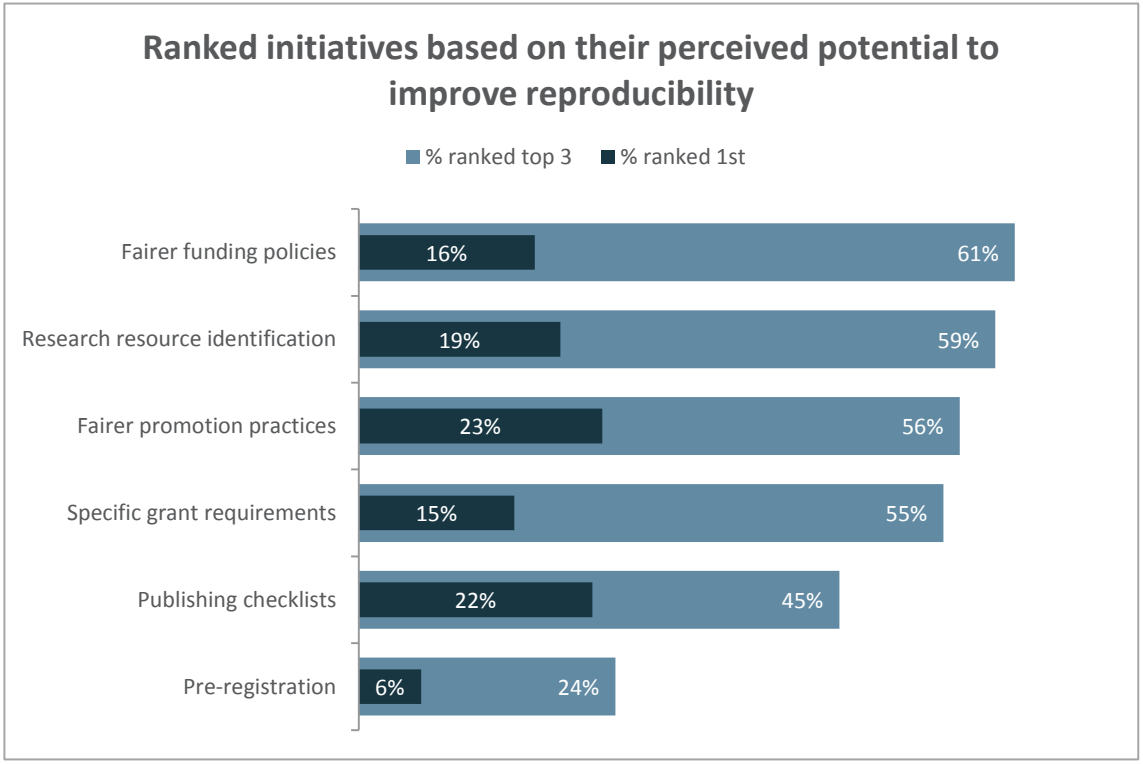
- The figure below illustrates the level of agreement that respondents had with three different statements about the effectiveness of the reproducibility initiatives instigated by research institutions, funding agencies and publishers.
- Interestingly, a greater proportion of respondents agreed that publisher efforts with regards to reproducibility had helped with their work (57%) and had a positive effect on their field (57%) in comparison to both research institutions and funding agencies. However, there was also a greater proportion of respondents who felt that publishers should do more to encourage reproducibility in their field (75%).



Nature journal 'Reproducibility' efforts (Oct' 2017)

Initiatives perceived to have the greatest potential to improve reproducibility of research

- The figure below illustrates the initiatives that respondents ranked first (dark blue) as that with the greatest potential to improve reproducibility overlaying the proportion that the same initiative was ranked in the top 3 (light blue).
- Practices that support fairer promotion was ranked first by the largest proportion of respondents (23%) followed closely by publishing checklists. However, policies that support fairer funding was ranked most frequently in the top three. Together, this could indicate that either no one initiative is likely to solve the issues as perceived by respondents or that the real solution was not listed.



Nature journal 'Reproducibility' efforts (Oct' 2017)

Views on Nature journals reporting checklist

4.0

Awareness of the Nature journals reporting checklist

- In 2013, *Nature* and Nature research journals introduced a reporting checklist (bottom left) to "prompt authors to disclose technical and statistical information in their submissions, and to encourage referees to consider aspects important for research reproducibility".
- Of the Nature journal authors who responded to the survey, all were aware of the checklist, of which a quarter stated to having submitted over 3 manuscripts using the checklist.

Experimental design

1. Sample size

   Describe how sample size was determined.

   Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.

2. Data exclusions

   Describe any data exclusions.

   If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.

3. Replication

   Describe whether the experimental findings were reliably reproduced.

   For each experiment, note whether any attempts at replication failed OR state that all attempts at replication were successful.

4. Randomization

   Describe how samples/organisms/participants were allocated into experimental groups.

   Describe how samples were allocated to groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.

5. Blinding

   Describe whether the investigators were blinded to group allocation during data collection and/or analysis.

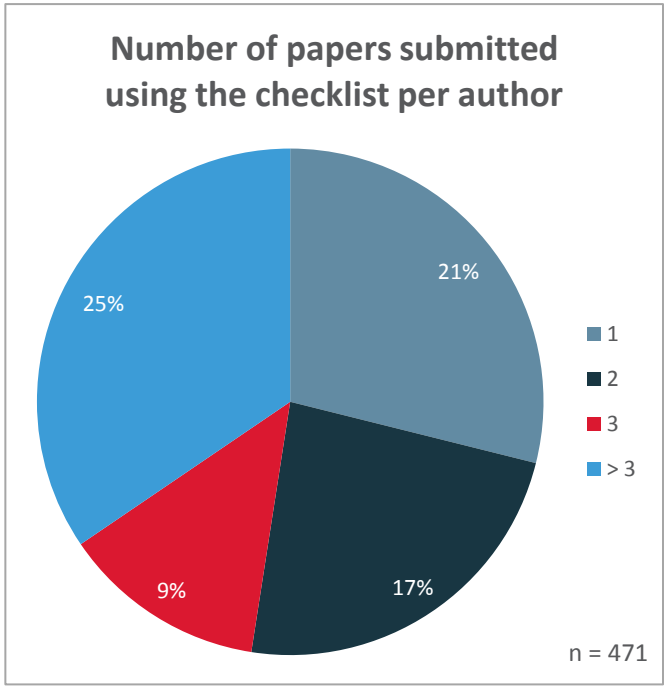
   Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.

Note: all studies involving animals and/or human research participants must disclose whether blinding and randomization were used.

6. Statistical parameters

   For all figures and tables that use statistical methods, confirm that the following items are present in relevant figure legends (or in the Methods section if additional space is needed).

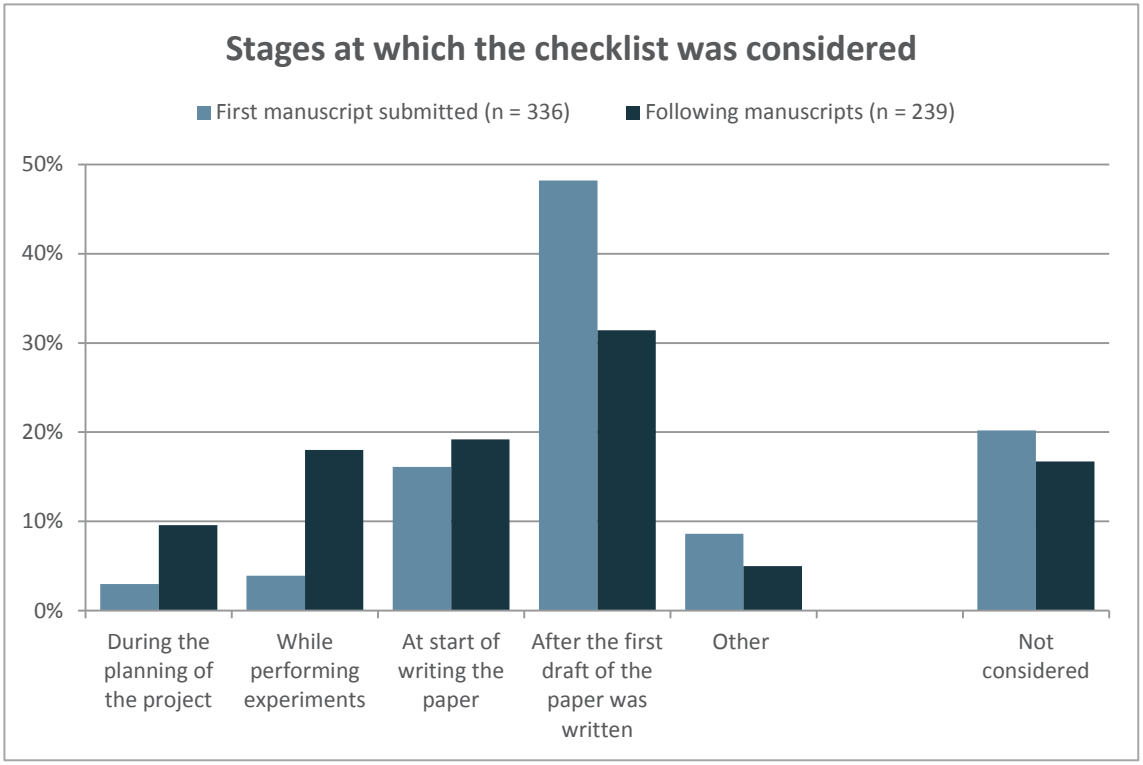
n/a	Confirmed
☐	☐ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement (animals, litters, cultures, etc.)
☐	☐ A description of how samples were collected, noting whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☐ A statement indicating how many times each experiment was replicated
☐	☐ The statistical test(s) used and whether they are one- or two-sided (note: only common tests should be described solely by name; more complex techniques should be described in the Methods section)



Nature journal 'Reproducibility' efforts (Oct' 2017)

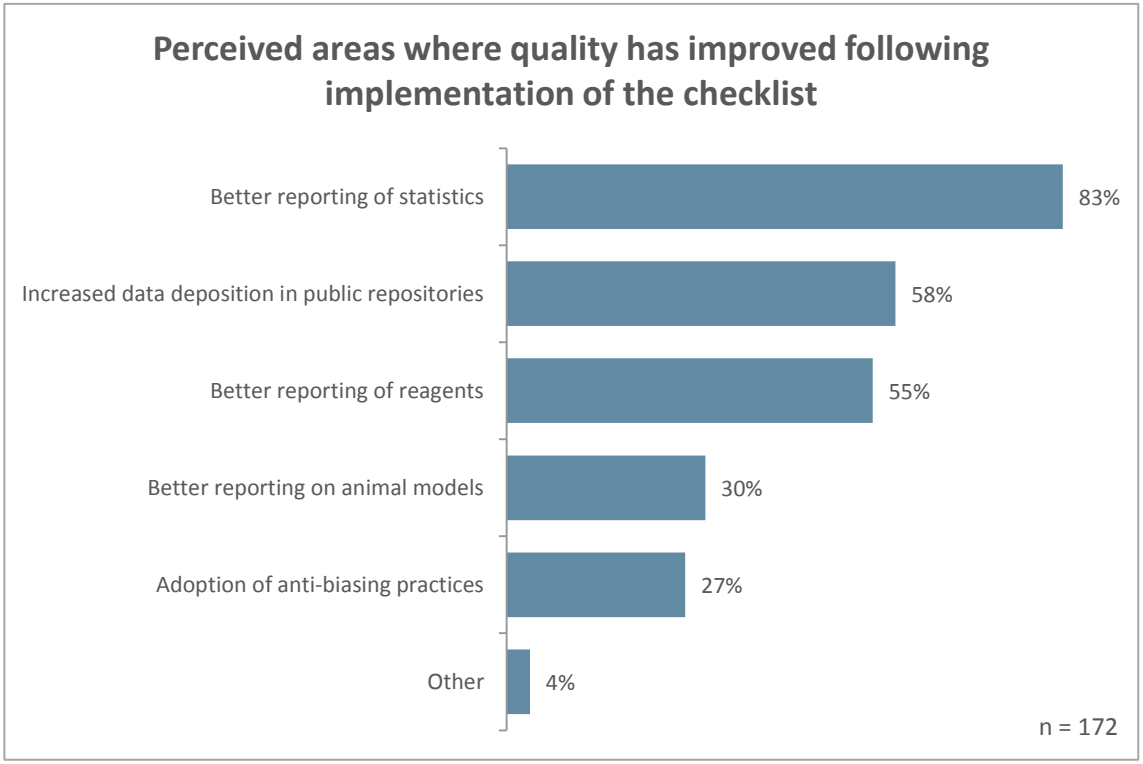
When is the checklist considered?

- The graph below illustrates the proportion of respondents indicating at which stage of their publication they considered the Nature journal checklist both the first time they submitted to a Nature journal and if they had, the second time as well.
- The first time a researcher submits to a Nature journal, the reporting checklist is predominantly considered after the first draft of the paper is written (48%) or not at all (20%).
- Encouragingly, following the first submission, consideration of the checklist earlier in the research process increases with 7% more respondents considering the checklist during the planning of the project and a further 14% considering while performing experiments. Importantly, the proportion not considering the checklist also decreases but only by 3%.



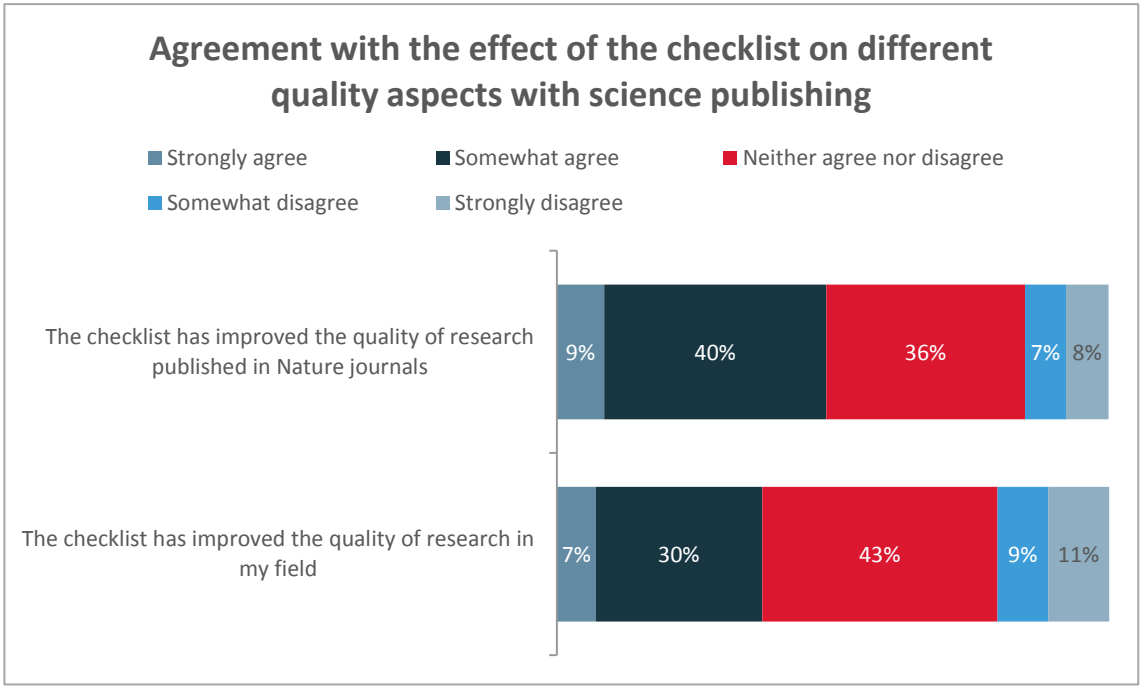
Perceived quality improvements due to checklist

- If respondents felt that the checklist had improved the quality of research at Nature journals (see slide 18), they were asked how they felt that the quality had improved.
- Of these, the one measure that the majority of panellists felt that the checklist had significantly improved was the reporting of statistics within the published papers (83%).
- Around a quarter of those who responded to this question felt that the checklist improved the adoption of anti-biasing practices which is important considering that selective reporting of results was felt to be a significant factor that led to problems with reproducibility. However, such a relatively small proportion may suggest that the checklist is missing or not addressing the key issue as seen by researchers.



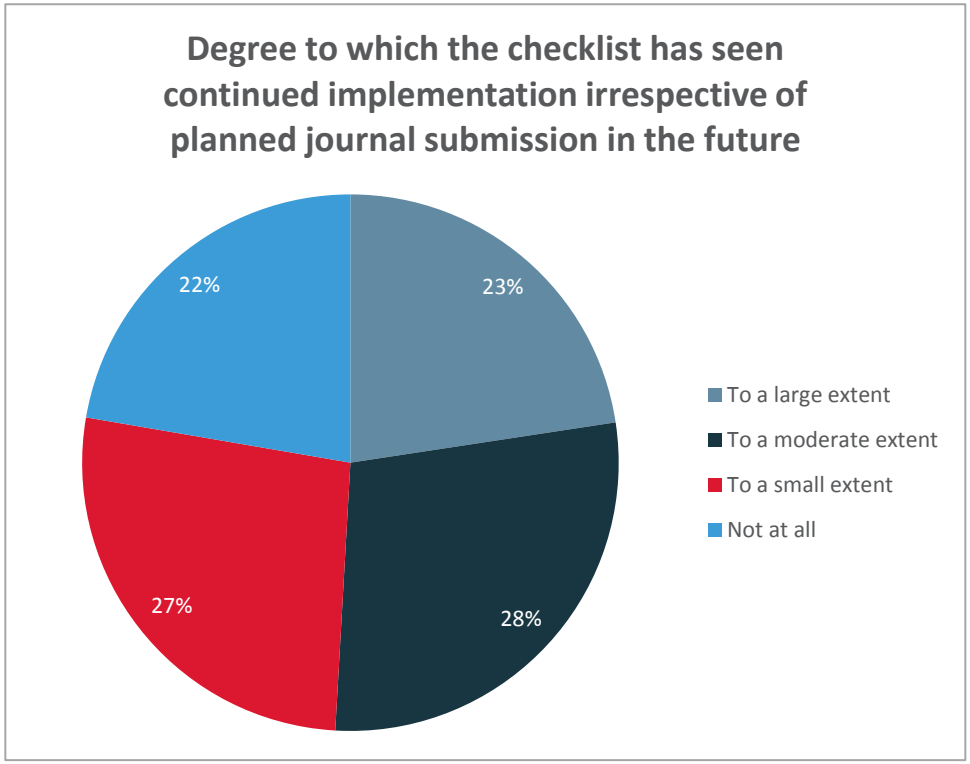
Effect of reporting checklist on overall quality

- At a top level, around a half of the respondents agreed at some level that the reporting checklist had improved the quality of research published in Nature journals.
- Similarly, 37% of respondents agreed that the checklist had also improved the quality of research in their field.
- It is of note however, that while there was not a lot of disagreement with these statements, there were large proportions of respondents who neither agreed nor disagreed with regards to the reporting checklists overall effectiveness. With particular regards to the effect on respondents' fields of interest, this may be representative of the perceived impact of the journal in the field or it may again reflect that the checklist is not addressing the core issues of reproducibility as authors see it (please see section 5.0 for further comment).



Effect of the checklist beyond Nature journal submission

- As a measure to determine the extent of the effect that the Nature journal reporting checklist is potentially having beyond submissions to Nature journal themselves, respondents were asked whether they had continued to implement the checklist irrespective of whether they had planned to submit again to a Nature journal in the future.
- Promisingly, 78% of respondents indicated that they had continued to implement the checklist at least to a small extent with their continued work with roughly a quarter stating they were using it to a large extent.



Further discussion

(text analysis)

5.0

Comments on the Nature journals reporting checklist

A key takeaway from the survey is the degree to which researchers believe there is an issue with regards to reproducibility of published findings (see slide 3). To add to this, the level of comprehensive text response and emails received in response to the survey indicates the high level of interest and opinion there is on the topic as a whole. As such, the following slides attempt to highlight some of the discussion. Full open text responses can be shared on request.

- Similar to the results about the overall effectiveness (see slide 18), there was positive sentiments towards the position of the Nature journals reporting checklist:

"The checklist is an important instrument. It forces authors and reviewers to check key aspects. / It won't solve all reproducibility issues, but should help identify more systematically weaker experimental design, inappropriate data analyses, context in which the experiment was carried out, etc." **– Research Scientist in the Earth and Environmental Sciences, Norway**

"I find the Nature checklist extremely helpful." **– Professor in the Life Sciences, Germany**

- However, a key issue with the checklist as perceived by a number of researchers is its ubiquitous and inflexible use across the journals when they felt that it should be adapted to the different needs states of their subject fields:

"I think the NPG checklists are right in spirit, but have little effect in practice. It's hard to make a checklist that is generically relevant to all the types of work submitted." **– Principal Investigator in the Life Sciences, UK**

"The nature checklist does not contain any requirements for source code, please add source code deposition and open-source requirements to the checklist." **– Professor in the Life Sciences, Canada**

"It is very field dependent. Much more of a problem with population and animal studies where statistics and sample sizes are so relevant. In my area, for example, when we complete the Nature checklist the vast majority of the questions are irrelevant. In my area inability to reproduce a study usually reflects sloppy science and poor practice by the researchers - this doesn't get picked up by a checklist for obvious reasons - it stems from factors such as pressure to publish and poor scientific ethics, which lead to the results of the study being interpreted incorrectly or too much." **– Professor in a multidisciplinary field, UK**

Survey missing the point

- For some researchers, the survey as a whole was a significant "let down" either addressing an issue they did not feel was relevant…:

"Reproducibility is a red herring. The aim of science is to identify mechanisms, not to satisfy fashionable buzzwords. Reproducibility may be a problem in areas on the margins of science, but in any good experimental area, either a study is built on or ignored. If it is built on, then of necessity people reproduce the original finding before proceeding. If enough people do not succeed, the study is ignored." - ***Email from Professor in the Life Sciences, US***

"Is reproducibility in a scientific sense really the key issue? Or is it honesty, transparency, self-criticism, and ethics in describing and reporting? The issues with areas from clinical trials to global climate change are not really about the narrow issue of reproducibility (in fact one can make a strong case the irreproducibility is an essential part of discovery-stage science, and that when it is gone, the field is sterile; surprises feed a growing field of science); rather, is it that scientists knowingly, or influenced by conscious or unconscious prejudice or ideology, fit or select their (and others') observations to fit their interpretations, rather than the other way around." – ***Email from Professor of Chemistry, US***

- … or on the opposite side of the fence, not addressing the perceived issues correctly in a manner to which they felt reflected their opinions on the topic. This highlights clearly the complex and multifaceted nature of the topic as well as the problems with trying to condense these issues into a concise all encompassing survey:

"There are real problems underlying the perceived "crisis of reproducibility" in science, but this survey does not give any room to discuss them … The problem, such as it exists, is with interpretation of work that is only weakly experimentally based and substitutes statistics for controls. Unfortunately, this is the kind of work that most often addresses questions that interest the general public. That is a valid type of inquiry but it is not the same thing as experimental, mechanistic science, and it is never as strongly reproducible or predictive as experimental science." – ***Email from Professor in the Life Sciences, US***

"I did the survey and feel that it asks the wrong questions and not tackling data reproducibility head on. The questions seem to elude that data sharing = reproducibility. Its not, data sharing is transparency, not reproducibility. There were limited adequate questions in the questionnaire that really capture how to improve reproducibility." – ***Email from Professor of Medicine, US***

"Most of these efforts would have little impact on reproducibility and distract from the central issue, which is the selective reporting of data to fit a pre-supposed model. This problem is widespread and not addressed by journals even when it is clearly evident." – ***Open text response from Laboratory Director in the Life Sciences, US***

"You didn't ask anything about the main source of problem in my area, which is that the scientists are unable to design and interpret the experiments properly. Therefore, it is not that they did not do enough replicates, provide the data etc., it is that they did not know how to interpret the data, did not do the correct experiments, or over interpreted the results. The same for reviewers, so that the poor experiments got published." – ***Open text response from Professor in the Life Sciences, UK***

Nature journal 'Reproducibility' efforts (Oct' 2017)

Different concepts and the cultural shift

- Many respondents wrote in to either discuss potential solutions and concepts to counter irreproducible research while other felt that singular ideas didn't go far enough and entire shifts in research culture was required to move the quality of data forward:

Reproducibility requires money and the funders need to provide that money for the research to be done. Funders only want to fund novel research, not something that has already been done. I think this is the dominant driver. If funders provided money specific for reproducing results and high tier journals, like nature, accepted publication of these kind of results, reproducibility would not be a problem.

"For several years now, we use the Supporting Information to provide the results from all of our replicate experiments, including what minority responses are observed, if they are observed, that may differ from the majority responses. This has been helpful to me, as the claims one can make in cases where minority responses are observed in addition to the majority response have become much broader." **– Professor in the Physical Sciences, US**

"Reviewers do not always pick up on issues with data (for a number of reasons). As an extra layer, it might be worthwhile looking more into how Journals can also scrutinize manuscripts/data. Editors should do this, but I think that rarely happens. Software can do part, but using some sort of data-assessors, alike copy-editors, might help." **– Principal Investigator in Life Sciences, UK**

"It comes across, frankly, as a sick joke that publishers are adding statistical checklists to the requirements for manuscript submission, while at the same time they are undermining the whole basis of scientific scholarship by limiting the number of references that can be cited in a manuscript. References are limited -- either by explicit format guidelines or by stringent character/word count limits. Authors cannot afford to cite primary research papers because they don't have enough citations to include them all. They cite reviews instead. The original research papers are not read, and are not even expected to be read. This hurts the careers of the people who wrote those original research papers, and it hurts the journals that published them. But the increasingly common practice of replacing first-hand with second-hand or third-hand citations like this also means that a lot of the references are simply WRONG. And no one cares -- the number of mis-citations is just too high. But it's not just about credit. Authors can assume that their new results are relevant to something, when in fact they (or the review writer whose review they are citing) have simply misinterpreted what was in the original paper. As long as authors don't ever read the original papers reporting relevant studies any more, and there is no penalty for lousy scholarship, people will be wrong a lot. This kind of ignorance is a major source of "irreproducibility"." **– Professor in the Life Sciences, US**

"In our research field a scary number of reports in highly rated journals including Nature are not reproducible. We made the experience that careful studies that uncover the flaws in such publications and provide solid, reproducible sets of data that correct these reports usually encounter a stunning disinterest by the same journals. Since jobs for young scientists are awarded according to their publication record in top journals, this non-reward for solid, reproducible and often somewhat slower science is an important element in a vicious circle promoting sloppy, non-reproducible and sometimes even fraudulent ´science´." **– Professor in the Life Sciences, Germany**

"We do such a poor job of this (creating reproducible research) that we need a cultural shift. Both bottom up, so it is a standard part of all graduate students' curriculum and top-down, so departmental heads and funders take a lead in encouraging good practice." **– PI in the Life Sciences, UK**

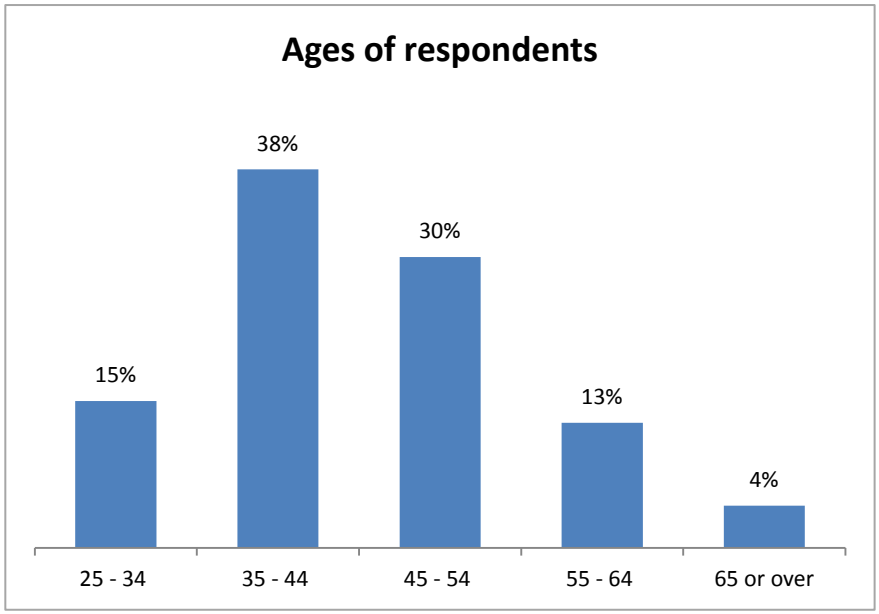
Nature journal 'Reproducibility' efforts (Oct' 2017)

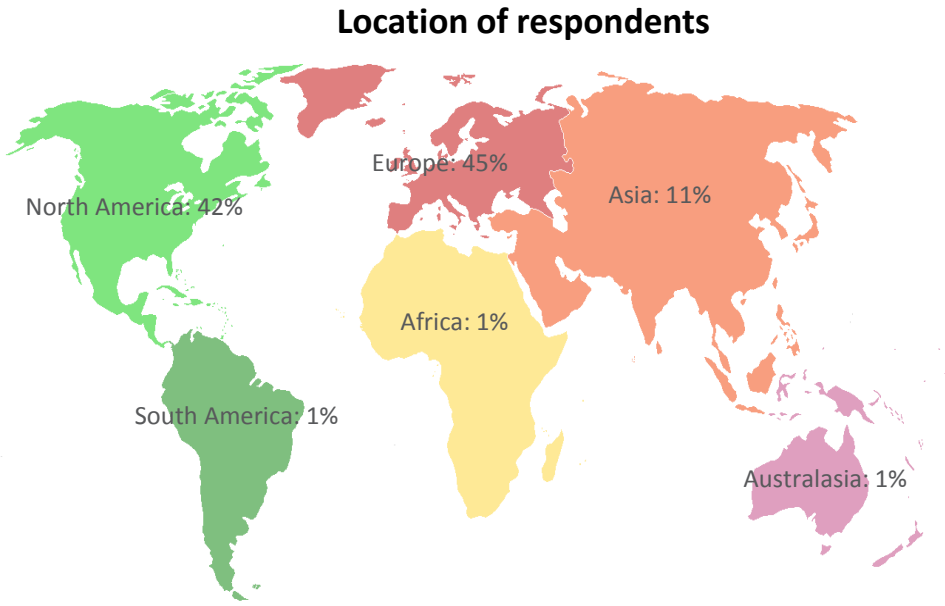
Appendix

6.0

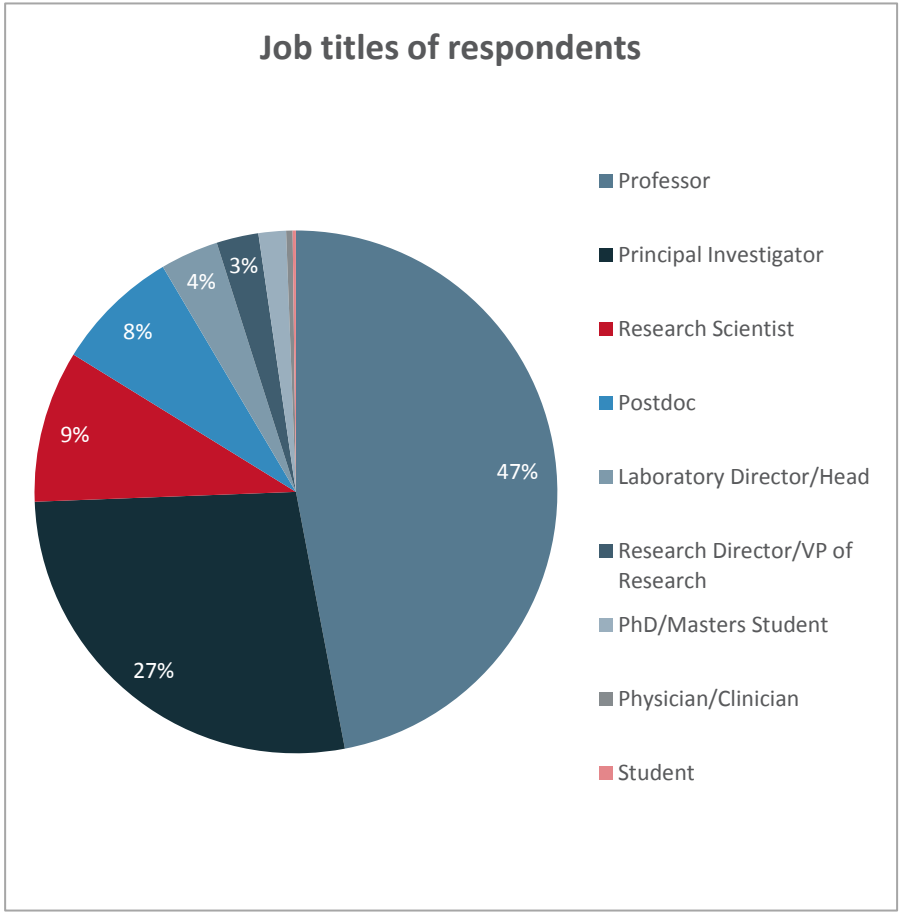
Demographics

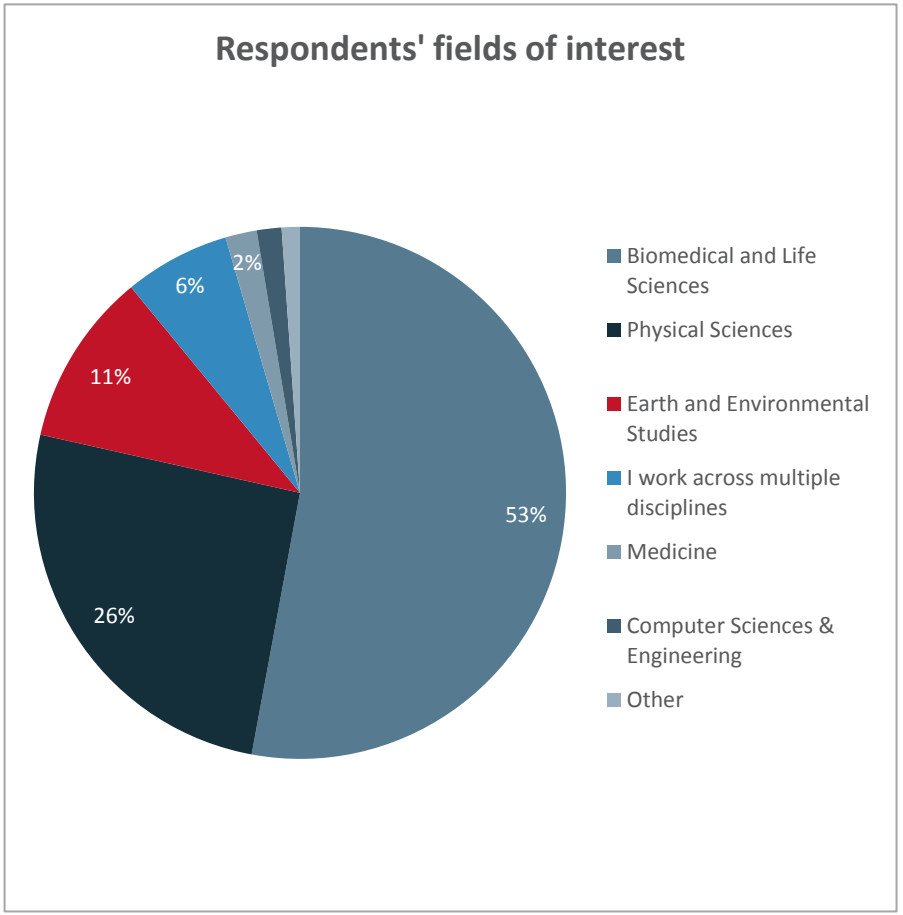
- The following slides provide information on the demographic breakdown of all respondents to the survey:





Professional Status





Nature journal 'Reproducibility' efforts (Oct' 2017)

The story behind the image

How chameleons change colour

Chameleons are well know for their potential to change colour but recent research on panther chameleons is the first to find two layers of crystal containing cells, each with a potentially different purpose. Researchers from the University of Geneva have speculated that the deeper crystal containing cells may help with the regulation of temperature, whilst the more superficial layer of colour changing cells could be responsible for camouflage or mating displays.