

Attitudes to sanctions for serious research misconduct: results from the FAIRS Delphi survey

version 1.6, 30th June 2025

Dorothy V. M. Bishop
University of Oxford
dorothy.bishop@psy.ox.ac.uk
ORCID: [0000-0002-2448-4033](https://orcid.org/0000-0002-2448-4033)

Abstract

Background: Research fraud is often seen as a rare event, but evidence from self-report surveys indicates that fabrication and falsification of data are common enough to be a problem. This study assessed attitudes to serious research misconduct, contrasting views of self-appointed research "sleuths" and research integrity officers (RIOs).

Methods: Respondents completed a two-round Delphi survey, rating agreement with statements about prevalence, responses to, and consequences of serious research misconduct. In Round 1, there were 95 respondents (37 sleuths, 33 RIOs and 25 other). For Round 2, there were 79 respondents (36 sleuths, 22 RIOs and 21 other).

Results: Respondents agreed on the damaging impact of academic incentive structures on research integrity, and the importance of agencies working together to address serious research misconduct and preserve the academic record. There was polarization of views between sleuths and RIOs on the frequency of serious research misconduct, the adequacy of institutional responses, and the suitability of self-regulation by institutions.

Conclusion: Sleuths and RIOs operate in information silos. They could benefit from working together, but this will only be possible if trust is restored. Given dissatisfaction with institutional self-regulation, drawbacks and benefits of alternative regulatory models need to be evaluated.

(196 words; limit 200)

Keywords: research integrity, fraud, self-regulation, Delphi method, research sleuths

Introduction

The past decade has seen a growth of articles about Questionable Research Practices (QRPs), but less attention has been given to research fraud. Although the lines between QRPs and fraud can be blurred (Bouter, 2023, Gopalakrishna et al., 2023), it is generally agreed that falsification, fabrication and plagiarism go well beyond "questionable" activities to correspond to serious research misconduct.

In responding to fraud, there is a sense of tension between two constituencies. On the one hand, there is a body of volunteer "sleuths" who report on fraudulent work and try to get it retracted. Most sources of information about sleuth activities are scattered and are not in peer-reviewed papers, with the bulk of evidence reported on the post-publication peer review website PubPeer (Couzin-Frankel, 2015). On the other hand, we have what may be termed "the establishment": publishers, academic institutions and funders who are seen as responsible for dealing with cases of fraud. Unlike sleuths, who can rapidly respond to apparent cases of fraud, the establishment has to ensure due process is followed, that malicious or trivial complaints are distinguished from plausible accusations of misconduct, that accused and accusers get a fair hearing, and that any actions they take are properly regulated.

Concerns about research fraud are international, but there is no agreed process for dealing with cases. A report by the National Health and Medical Research Council of Australia (2023) surveyed practices in nine different international jurisdictions and found a wide range of regulatory frameworks.

Materials and methods

A two-stage Delphi survey was used to gather both quantitative and qualitative responses to a set of statements covering the topic of research fraud. The Delphi method (Hasson, Keeney & McKenna, 2000) has the advantage that it can bring together a wide range of opinions without it being necessary for individuals to meet in person. Participants completed an online survey where they rated strength of agreement with statements, and could add textual comments to justify their position. After one round, a group of moderators assembled the responses and each panel member received feedback showing how their responses compared with those of (anonymized) other respondents. The moderators could, at this point modify problematic items. Panel members were then sent a summary of Round 1 results, plus revised survey items for Round 2, with a document explaining the rationale for any changes. A protocol for the survey can be found in Supplement 1.

Ethics approval for this study was obtained from University of Oxford Central Research Ethics Committee (CUREC), application MS IDREC 750906.

Participants

Advisory board

A three-member Advisory board advised on selection of moderators and the expert panel, selection of the statements to be evaluated, and criteria for deciding on which items to retain, reject or modify. The Advisory Board included an expert on Delphi methodology, a senior

research scientist who had published a book on irreproducible science, and an expert in research integrity investigations (see Supplement 2).

Moderators

The moderators (see Supplement 2) were responsible for running an initial exercise to determine which statements were included in the Delphi, for selecting the expert panel, and for deciding which statements should be retained, dropped or modified between Rounds 1 and 2 of the Delphi. They did not themselves participate in rating of any statements, but they were experts in the area and able to evaluate qualitative responses by panel members.

Selection of panel members

The goal was to include panel members to cover a range of areas of expertise, subdivided into two groups that may be expected to respond differently to the survey:

Group S (for Sleuths) aimed to include victims of research fraud, whistleblowers, and sleuths/forensic metascientists

Group R (for Research Integrity Officer) aimed to include integrity staff of academic publishers, university integrity officers, staff from research funders, lawyers, and government representatives.

We relied on individuals' self-descriptions to assign them to these categories. In practice, around one third of respondents did not align specifically with either group S or group R. These included some who had experience in both roles: detecting research fraud and sitting on institutional panels, as well as others with an academic or policy-based interest in the topic. An "Other" group (O) was created for those who did not clearly belong in Group S or Group R.

A core sample of potential panelists had been identified when planning for the related academic meeting, and we used snowball sampling to extend this, writing to these individuals with our desired criteria for respondents, and asking them to nominate suitable people.

Individuals were invited to express interest by completing an online form. Recruitment was monitored across stakeholder groups and geographical location to improve recruitment for groups in which recruitment numbers were low by targeting potential participants in these groups. The aim was to include at least 30 expert panel members in each of Groups S and R, but additional people who expressed a wish to take part were accepted on the panel, and allocated to Group O.

Selection of statements for inclusion

A preliminary pool of items was generated using three approaches. First, statements were devised to cover important and possibly controversial issues that were discussed in academic articles and institutional and governmental reports on this topic. Second, the website <https://lmarena.ai/> was used to gather AI-generated responses to the prompt: "Create statements for a Delphi project on the topic of how institutions respond to academic fraud." The AI-generated items came from two models: gpt-4o-2024-05-13 or engine-test. Third, further items were created to cover additional aspects that were deemed potentially relevant on the basis of discussion with the Moderators. The preliminary item pool is available as an Appendix to the protocol (Supplement 1).

This pool was taken as a starting point for discussion with the Moderators, who discussed which items should be included or excluded and recommended improvements to the wording. There was some debate around how to refer to the central concern of the survey. Those working in the research integrity field preferred "research misconduct" to "research fraud", in part because the latter term has legal implications. On the other hand, research misconduct covers a much broader scope than research fraud, which is typically limited to plagiarism, fabrication and falsification. It was agreed to use the term "serious research misconduct" in the survey, and to define it at the outset as referring to plagiarism, fabrication and falsification.

Procedure

Round 1

People who expressed interest in joining the panel were asked to complete a short form asking for basic background and demographic information, to allow categorization based on domain(s) of expertise (Supplement 3). All those who volunteered were invited to join the panel. They received a personalized email providing a link to the survey, which was implemented online using JISC online surveys. The full text of the survey is available in Supplement 4. Respondents were told that their identity would not be disclosed to other panelists or otherwise made public and were advised not to include any identifying information in their responses.

Moderator evaluation

The moderators met to scrutinise the pattern of responses and the text comments. In a typical Delphi survey, items are rejected if agreement is poor. This was not appropriate for this study, where we aimed to identify items where Group R and Group S agreed or disagreed. In addition, free text responses were scrutinized to consider whether it is likely that stronger consensus might be obtained by rewording. At this point, some statements were dropped or reworded. In addition, a selection of relevant and constructive comments from panel members were incorporated in the explanation of statements that went forward to Round 2 to provide context. A written justification for any changes to the item pool is available in Supplement 5.

Round 2

Panel members were presented with an individualized report showing both their own scores and the anonymized distribution of scores from other panel members, together with a sample of anonymized free text comments that illustrated different viewpoints. Panel members were asked to complete the revised survey in light of this feedback. The Round 2 survey again included a free-text box for each statement so that respondents could elaborate on their responses.

Results

Departures from protocol

There were three departures from the planned protocol. First, rather than requiring all respondents to identify with Group R or Group S, we included an Other group. Results from these panel members are included in the summary tables, but were not used to categorize items in terms of agreement.

Second, a few items departed from the 5-point scale structure and were multiple choice. This allowed us to capture more qualitative aspects of preferences for some items. A couple of items in Round 1 involved rank ordering options rather than rating them, but this confused some respondents and these items were converted to the rating format in Round 2.

The third departure concerned criteria for categorizing items as reaching agreement, or showing polarized views. The original plan was to use a cutoff of 80% endorsement as evidence of agreement; this was relaxed to 70% to allow a higher proportion of items to be included in the results. The original plan of reporting mean and median ratings was dropped on advice of the Advisory Board; the percentage agreement scores are easier to understand and more suited for ordinal data.

Panelists

For Round 1, 116 panelists signed up to take part in the survey, and completed surveys were received from 95 of them. There were 64 men and 31 women, and the bulk of respondents were from Europe (N = 67) or North America (N = 18), with 4 from Africa, 1 from Asia, 3 from Australasia, and 2 from South America.

All of those who had volunteered for Round 1 were invited to take part in Round 2, including those who had not completed Round 1. 97 panelists signed up to take part in Round 2, and completed surveys were received from 79 of them (one of whom had not completed Round 1). There were 53 men and 26 women. Geographic distribution was 55 from Europe, 14 from North America, 3 from Africa, 2 from Asia, 3 from Australasia, and 2 from South America.

In terms of subgroup identity, in Round 1, participants had been asked:

We anticipate that patterns of responses may be different for those who are primarily involved in raising concerns about serious research misconduct (e.g. sleuths, whistleblowers, victims of research fraud) and those whose job it is to respond to reports of such concerns (e.g. research integrity officers, those devising policies for research integrity). Could you please specify here which of those groups you identify with most strongly - or whether neither applies to you.

- *Group 1: Sleuths, etc*
- *Group 2: Research integrity officers, etc*
- *Group 3: Neither of the above*

The proportion who responded "Neither" was higher than expected, so In Round 2, the question was asked again, but with the additional prompt:

"Some panel members had roles both as sleuths and in research integrity or other roles. If that is the case for you, please select the role you feel most strongly affiliated with, rather than selecting 'Neither'."

In addition, respondents who answered "Neither" were given the opportunity to specify their role.

We had anticipated that this additional instruction would encourage more individuals to select group S or R, but it did not achieve this effect, as respondents were as likely to move out of these groups as into them. In terms of those who dropped out (i.e. Missing) in Round 2, there was no obvious pattern, with similar levels of dropout in all three subgroups.

Table 1: Self Identification in Rounds 1 and 2

Round 1	Sleuth (S)	Round 2		
		Research Integrity Officer (R)	Neither /Other (O)	Missing
Sleuth (S)	29	0	2	6
Research Integrity Officer (R)	1	20	5	7
Neither /Other (O)	5	2	14	4
Missing	1	0	0	0

Those who responded Neither included journal editors, academic researchers, and those with a research interest in research integrity.

Quantitative data: aggregated responses by subgroup and level of agreement.

For the current report, the focus is on Round 2 results, with items grouped into those where there was agreement between groups S and R, and those where there was polarization of views.

The full set of quantitative data, showing percentages of endorsement of all statements in Rounds 1 and 2 is shown in Supplement 6. As noted above, some respondents changed their subgroup status between rounds: In all tables, the Round 2 categorization of subgroup was used for both Round 1 and Round 2 ratings, so that consistency of responses across Rounds could be evaluated.

Tables of aggregated responses show percentages endorsing a statement. Most sub-items were rated on a 5-point Likert scale, and the percentages show those responding 4-5. A few items had multiple-choice options rather than a scale, and for these the percentage endorsing each option is shown.

Data for subgroups and the whole sample are shown colour coded, where percentage of panelists agreeing with a statement is presented as follows:

- blue denotes > 70% agree with the statement
- pale blue denotes > 50-69% agree with the statement

- pink denotes 30-49% agree with the statement
- red denotes < 30% agree with the statement

This coding was done to make it easier to visualize cases of agreement or polarization. Items were coded as showing agreement between Groups S and R if the responses went in the same direction (i.e. all either red/pink or blue/pale blue) and at least one group had at 70% or more agreement (blue) or less than 30% agreement (red). This criterion excluded items where there was not much agreement within groups (e.g. if both groups were coded as only mild agreement).

Items were coded as polarized if the percentages agreeing went in opposite directions, and one group had at least 70% agreement (blue) or less than 30% agreement (red).

The organization of the Results section does not follow the serial order of survey items, but rather has items (and subitems) grouped according to whether they indicated agreement across subgroups of panelists, or whether responses were polarized. Responses are grouped into themes, identified by alphabetic letters.

Qualitative data: free text comments

The full sets of free text comments for Rounds 1 (Supplement 7) and 2 (Supplement 8) are available online. Selected comments are included in the current report with the quantitative data for some items, to give insights into panelists' justifications for their responses. Spelling errors have been corrected in the comments, and omitted material is denoted by Free text comments are identified by a number denoting the Round (1 or 2) and the item (after decimal point), followed by the Group of the respondent with a random number ID. At the request of one participant, who had concerns about anonymization, the random number given with each comment varies from item to item so it does not identify a specific individual.

Propositions with general agreement regardless of group

A. Serious research misconduct is an important issue, with negative impacts on several constituencies. Tackling it is important to preserve the integrity of the academic record.

This topic was covered by Item 3 in Round 2 of the survey:

Table 2. Item 3: "How harmful are the impacts of serious research misconduct to different segments of society? Please code as 1 (low harm) to 5 (strong harm)."

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
a) Consumers of research findings, e.g. patients whose treatment is informed by medical research or policy-makers who depend on research findings	77.8	100	100	89.9
b) Other researchers who try to build on fraudulent findings	80.6	95.5	71.4	82.3
c) Other researchers who are based in the same lab, or collaborate with, a fraudulent researcher	80.6	77.3	61.9	74.7
d) Funders, whose funds are wasted	69.4	77.3	42.9	64.6

The strongest endorsement overall was for research with practical impacts for users. A sample comment:

- > 2.3_S2) There is evidence that serious misconduct in papers delay treatment and development of new treatment

The relatively low rating for funders was surprising; it is possible that this will depend on the kind of funder one has in mind, as illustrated by this comment:

- > 2.3_S4) I have rated the funders lower for consequences, as they often have other investments to mitigate any impact from research misconduct, they are usually left out of being held accountable, and yet have a relatively high importance to those publishing through (monetary) pressure and power...

Strong agreement is also evident in responses to option (a) for item 4; preservation of an accurate academic record was a high priority for all.

Table 3 Item 4a. "In responding to serious research misconduct, several goals may be considered. Please rate how important each of these is, from 1 (unimportant) to 5 (very important)."

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
a) To correct the academic record	86.1	86.4	71.4	82.3

B. The incentive structure of academia, which evaluates people in terms of their publications, is a driver of serious research misconduct.

Item 6 specifically asked about factors encouraging researchers to commit serious research misconduct, and all groups thought the 'publish or perish' culture was a serious problem.

Table 4. Item 6c: "What is the impact of these factors in encouraging researchers to commit serious research misconduct? Please rate from 1 (little impact) to 5 (large impact)."

Percentage endorsing (4-5)

Subitem	1_S	2_R	3_O	All
c) 'Publish or perish' culture and incentive structure in research (e.g. progression, promotion, recognition)	91.7	95.5	81.0	89.9

One respondent noted a relevant report:

- > 2.6_R17) We know the incentive structure is to blame: <https://www.vitae.ac.uk/vitae-publications/research-integrity-a-landscape-study>

Another noted that there had been formal modeling of impact of incentives:

- > 2.6_S9) This is a paper that deals with rational choice theory (Sattler et al, 2012).

However, others cautioned about drawing generalizations. For instance:

- > 2.6_S4) ... Well-established researchers probably have little concern about their jobs; brand-new researchers, researchers on a contract, researchers who do not have tenure presumably care a lot about building their publication record.
- > 2.6_S6) Huge variations between countries. There are (still) places where it is practically impossible for a researcher to lose their jobs and serious research misconduct does very much exist there.

In Item 12a, there was general agreement that changing criteria for hiring/promotion/funding should help reduce research misconduct.

Table 5. Item 12a: "Given that we have finite resources, which solutions to serious research misconduct should be prioritised in funding?"

Percentage endorsing (4-5)				
Subitem	1_S	2_R	3_O	All
a) Changing criteria for hiring/promotion/funding away from publish/perish model, so that researchers won't be motivated to commit fraud	66.7	81.8	61.9	69.6

Nevertheless, some respondents thought this aspect had been exaggerated;

- > 2.12_R13) The publish or perish argument is not what I see as part of the investigations I have been assisting with; it is often about behaviours, lack of respect for students and failure to recognise contributions combined with arrogance around interpretation of data.

C. Need for support for individuals/organizations who have expertise in detection of serious research misconduct

This was an option from item 12, which asked about prioritising solutions to serious research misconduct. As might be expected, the strongest support came from sleuths, but overall, over 70% of respondents endorsed this.

Table 6. Item 12e: "Given that we have finite resources, which solutions to serious research misconduct should be prioritised in funding?"

Percentage endorsing (4-5)

Subitem	1_S	2_R	3_O	All
e) Support for individuals/organisations who have expertise in detection of serious research misconduct	72.2	59.1	81.0	70.9

Although it was frequently endorsed, there were not many comments on this option. This one noted that it would be relatively easy to adopt:

- > 2.12_S7) Funding for the work and legal protection of sleuths would make a real difference much faster than the other options. It doesn't require a culture change or big reforms on the level of institutions....

D. Vexatious/trivial accusations are not a major factor hindering investigations of misconduct

These data come from Round 1, item 5, which asked about factors that may hinder academic institutions' response to serious research misconduct. Because so few respondents endorsed this subitem, it was dropped from Round 2.

Table 7. Item 5f: "Various factors may hinder academic institutions' response to serious research misconduct. Please rate the following from 1 (not much of a hindrance) to 5 (substantial hindrance)."

Percentage endorsing (4-5)				
Subitem (ROUND 1)	1_S	2_R	3_O	All
f) Large number of vexatious/trivial accusations of misconduct	13.5	25.0	12.5	17.2

E. In tackling serious research misconduct, funders, publishers and research institutions need to work together.

Three subitems in Item 13 explicitly asked about sharing of information between publishers and institutions, and all were strongly endorsed.

Table 8: Item 13c, e and f: "Publishers and their journal editors are responsible for ensuring the literature is decontaminated from erroneous work promptly. Please rate the following statements in accordance with your views on how this should work, from 1 = strongly disagree to 5 = strongly agree"

Percentage endorsing (4-5)				
Subitem	1_S	2_R	3_O	All
c) Publishers and institutions should work together to facilitate sharing of key information when serious research misconduct is suspected	94.4	86.4	90.5	91.1
e) When an institutional investigation is completed, the institution should directly approach the publisher of articles affected by severe research misconduct and request retraction	97.2	90.9	100	96.2
f) When a publisher finds the same author repeatedly flagged for erroneous material, they should communicate this to research integrity officer at the author's institution	80.6	95.5	90.5	87.3

F. A low probability of being detected/reported encourages serious research misconduct

Item 6 asked about the impact of factors encouraging researchers to commit serious research misconduct. Option (a) met the criteria for agreement, though there was a 30% percentage point difference between sleuths and RI officers.

Table 9. Item 6a. "What is the impact of these factors in encouraging researchers to commit serious research misconduct? Please rate from 1 (little impact) to 5 (large impact)."

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
a) Low probability of being detected and/or reported	86.1	54.5	61.9	70.9

G. Post-publication peer review via PubPeer is on balance a useful source of information about serious research misconduct

Item 7 asked about various sources of post-publication peer review. Only PubPeer achieved an overall positive rating, with only sleuths having more than 70% agreement.

Table 10. Item 7a: "Please rate the impact of the following from 1 = strongly negative to 5 = strongly positive, in drawing attention to serious academic misconduct"

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
a) PubPeer	80.6	63.6	52.4	68.4

Some responses from Group R noted problems with anonymity of comments (which is optional with PubPeer), as exemplified by this comment:

- > 2.7_R12) whilst we recognise PubPeer is used by a large number of academics, we do not find it particularly helpful in investigating misconduct as anonymous comments could be malicious, and if they are genuine the anonymity prevents us from following up with the complainant or gathering more evidence.

H. Conflict of interest for institutions investigating their own researchers is a hindrance in dealing with serious research misconduct

This item met our definition for agreement between groups, although the endorsement was considerably stronger for sleuths than for other groups.

Table 11. Item 5b. "Various factors may hinder academic institutions' response to serious research misconduct. Please rate the following from 1 (not much of a hindrance) to 5 (substantial hindrance)."

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
b) Conflict of interest for institutions investigating their own researchers	91.7	52.4	52.4	70.5

I. There should be sharing of information between employers, funders and publishers to support investigations of serious research misconduct

Item 11 consisted of a single statement, with multiple choice options indicating level of agreement.

Table 12. Item 11: "Employers, funders and publishers of research should be legally required to share information to support investigations of serious research misconduct". (multiple choice options).

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
Moderately/strongly agree	88.9	81.8	100	89.9

Although most agreed, negative consequences were noted:

- > 2.11_S5) Another tough call. How can this be protected against being weaponized by malicious actors?
- > 2.11_R12) ... Sharing of information to support investigations already takes place in some countries, but this is carefully supported by regulation and the law as well as with trust that there will not be inappropriate breaches of information. If trust in stewardship of information decreases in a nation-state then sharing of information may no longer be either possible or desirable and we also may end up with people too scared to enter a scientific career in the first place...
- > 2.11_O19) As a former funder, I know the frustration of not being provided with the information necessary to conduct an investigation or not being appraised of an institutional investigation instigated by us. That said, great care needs to be taken on how shared information is handled to ensure its security.

Relatedly, in item 10, over 80% of all respondents agreed that prospective employers should check with previous employers if there have been any investigations into serious research misconduct (summing options a and b). However, for this item there was disagreement about whether sharing should occur regardless of outcome (a), or only if research misconduct was found (b).

Table 13. Item 10. "Prospective employers should undertake rigorous due diligence and, as far as possible, check with previous employers to ask if there have been any investigations into serious research misconduct" (multiple choice options).

Percentages selecting each option

Options	1_S	2_R	3_O	All
a) Agree regardless of outcome	50.0	22.7	19.0	34.2
b) Agree if research misconduct found	33.3	59.1	66.7	49.4
c) Neutral	11.1	13.6	14.3	12.7
d) Disagree	5.6	4.5	0.0	3.8

For both items 10 and 11, commenters noted some practical and ethical difficulties in implementation, indicating that discussion should focus more on how to achieve effective, ethical and legal information-sharing, rather than whether information-sharing should occur. One respondent noted that there is precedent for sharing in the UK General Medical Council.

- > 2.10_S1) If there were a national body that investigated research misconduct that body could keep a record of "findings of misconduct" against a researcher. Anyone can search the General Medical Council's website for a doctor's record to see whether he/she has any GMC misconduct findings - it gives a summary of the findings and sanctions imposed, which range from conditions imposed to removal from the register.
- > 2.10_S7) In addition, applicants should be asked to certify that they have not been subject to investigations that found them "guilty". If they lie, then there is a paper trail, which makes dismissal easier. I suspect that privacy laws will make the idea of cross-institutional collaboration difficult (and for good reason).
- > 2.10_R18) ... On the one hand, we want to ease up the tensions around misconduct and foster mutual trust so that researchers would be more open about questionable practices, share their experiences and try to change the practices for the better together. On the other hand, there is the tendency for greater sanctions, publishing serious cases and tracking the employment history of researchers. The process of change should begin with open dialogue and honest discussions, followed by sanctions if the research community agrees to it and deems it necessary.... As a side note, perhaps due the current political environment, this type of tracking has strongly negative connotations.
- > 2.10_R19) ... sharing information that investigation took place yet didn't conclusively show that misconduct had occurred would be highly problematic and could unduly stigmatize a researcher. A tricky question would be how to handle ongoing investigations, where the result is still unclear.
- > 2.10_R17) ... I would like ... that all institutions make part of their application forms a question that seeks to understand whether there are any ongoing investigations, whether there have been any upheld allegations of research misconduct and if so, what was the type of misconduct.

J. A regulatory agency independent of Government is a potentially useful model for addressing serious research misconduct

This was the only option from item 9 that had overall approval, though only Sleuths achieved above 70% endorsement.

Table 14. Item 9d: "In an ideal world where resources are not an issue, which is the most suitable model/system for addressing serious research misconduct? Rate these options where 1 = least preferred and 5 = most preferred."

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
d) Regulatory agency independent of Government	75.0	59.1	47.6	63.3

K. There needs to be more protection for whistleblowers

This topic was the focus of item 14.

Table 15. Item 14a-c. "Please rate your agreement with the following statements about whistleblowers from 1 = strongly disagree to 5 = strongly agree."

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
a) There are disincentives for researchers to report serious research misconduct	86.1	63.6	61.9	73.4
b) Usually there is insufficient protection for whistleblowers who report serious research misconduct	86.1	63.6	76.2	77.2
c) Whistleblowers should have their identities protected, with confidential channels for reporting suspected serious research misconduct	86.1	72.7	71.4	78.5

Although respondents agreed with these points in general, they noted difficulties in real-world contexts where it may be unclear who was in the right. The idea of anonymity of whistleblowers was challenged by some on the grounds of natural justice:

- > 2.14_S1) "Whistleblowers should have their identities protected" - I agree up to a point. There has to be some consideration that a whistleblower may not be operating in good faith, and I think the person suspected of committing research misconduct needs to have the ability to robustly defend themselves. The criminal defense system has walked this tightrope for many years - protecting both the rights of the accused and the rights of the accuser... and it doesn't always work.
- > 2.14_R9) It is extremely difficult to deal with anonymous allegations, as the motivations of the Complainant have to be considered as part of the initial assessment.
- > 2.14_R11) ... If someone wants to make an allegation of misconduct via an institution that is not their own they do not have the same protections as if they were whistleblowing against their own employer - this can then leave whistleblowers really exposed to legal challenge should the researcher at that institution decide to sue them for defamation

L. There need to be meaningful sanctions for those committing serious research misconduct.

This was covered by two subitems of Item 15.

Table 16. Item 15: "Which of these practices should be options for institutions when serious research misconduct is confirmed. Please give your rating from 1 = strongly disagree to 5 = strongly agree".

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
c) Sanctions such as demotion or dismissal	91.7	72.7	81.0	83.5
d) Report on the investigation made public, with identities of those found to have committed serious research misconduct disclosed	86.1	63.6	71.4	75.9

M. Publishers should flag erroneous material promptly

This was covered in two options of item 13. The Sleuths gave particularly high endorsement for retracting erroneous material promptly.

Table 17. Item 13b and 13d. "Publishers and their journal editors are responsible for ensuring the literature is decontaminated from erroneous work promptly. Please rate the following statements in accordance with your views on how this should work, from 1 = strongly disagree to 5 = strongly agree".

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
b) Articles that contain serious errors that undermine confidence in the findings should be retracted promptly, without waiting for an institutional investigation	80.6	54.5	66.7	69.6
d) If conditions for retraction are not met, an expression of concern may be added to an article while an institution conducts an investigation	83.3	71.4	52.4	71.8

Some responses indicated disagreement regarding whether retractions imply misconduct:

- > 12.3_S10) Cannot be retracted before an official investigation; this is guilty until proven innocent
- > 2.13_R16) ...The second statement: in general, a factual error doesn't need to be investigated by a body, since it can be simply proven to be correct or not. Investigation attempts to establish factual basis to assess whether the erroneous conduct could or should be considered misconduct and whether the researcher is responsible (perhaps it was an honest mistake). In a similar way, if an author asks to rectify a mistake in a publication, it shouldn't mean that an investigation into potential misconduct should be initiated. There should be room for error checking and correcting outside of misconduct procedures. Making errors is not misconduct, it is quite human to make mistakes and err.

N. Those caught up in research misconduct investigations need protection

Although the survey mainly focused on what to do when serious research misconduct was proven, there were several points where respondents expressed concerns about the need to protect individuals caught up in investigations - either innocent individuals who were collaborating with someone who was the focus of investigation, or researchers who had an accusation against them that had not been proven.

There was good agreement regarding the first situation.

Table 18. Item 14d. "Please rate your agreement with the following statements about whistleblowers from 1 = strongly disagree to 5 = strongly agree".

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
d) It is important to be aware of and mitigate collateral damage that may be caused to other members of a research group if one member is found to have committed serious research misconduct	72.2	81.8	85.7	78.5

O. It is not justified to claim that investigations of serious research misconduct divert researchers and associated resources from more productive research work

This claim is sometimes made by those who think research integrity investigations have become too complex and unproductive (Caron, Lye, Bierer, & Barnes, 2024). However, responses to item 16 showed it did not receive general support from either Sleuths or RI officers.

Table 19. Item 16b: "Please rate your agreement with the following statements about unintended consequences/barriers to progress from 1 = strongly disagree to 5 = strongly agree".

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
b) Investigations of serious research misconduct divert researchers and associated resources from more productive research work	27.8	31.8	23.8	27.8

Propositions where views of sleuths and RI officers are polarized

P. Frequency of serious research misconduct

Item 2 explicitly asked respondents to choose between three options regarding frequency of research misconduct

Table 20. Item 2: "How common is the problem of serious research misconduct? (select one)"

Options	Percentage selecting each option			
	1_S	2_R	3_O	All
a) Serious research misconduct is rare relative to the amount of published research literature	11.1	40.9	38.1	26.6
b) Serious research misconduct is becoming more prevalent and starting to pose a threat to the research literature	16.7	36.4	33.3	26.6
c) Serious research misconduct is already common enough to pose a major threat to the research literature	72.2	22.7	28.6	46.8

The proportions of sleuths and RI officers endorsing option (c) was remarkably similar for Rounds 1 and 2. RI officers did not in general think serious research misconduct was frequent, although they thought it had important impact when it did occur. Some commented that sleuths had a distorted view because detecting fraud was a major focus of their activities. Sleuths, on the other hand, thought fraud was common but in most cases it went undetected, or, if detected, it was not adequately dealt with.

To some extent, the differences might reflect the impact of paper mills, which have injected a great deal of fraudulent material into the research literature (COPE & STM, 2022; Abalkina et al., 2025), but which are seldom the focus for research integrity investigations.

Nevertheless, the sleuths also referred to cases where clear evidence of fraud was ignored or covered up and where paper mills were not involved, and specific sources of evidence were cited (Berrío & Kalliokoski, 2024; Carlisle, 2021).

Here is a selection of relevant comments from Round 2 for this item:

- > 2.2_S1) Due to my own investigations I know of several fields that have been heavily targeted / affected by articles with evidence of serious research misconduct. There will be a very high chance that genuine researchers will use articles for reference that are completely fabricated....
- > 2.2_S7) ...Having worked in systematic reviews for 25y on similar biomedical topics it's clear to me that fabricated trials in my areas have become more prevalent, and they do affect the literature.
- > 2.2_R22) Talking about research misconduct has become more prevalent - but the number of proven allegations at my institution have not increased...
- > 2.2_R23) The issue we find is lack of reporting...
- > 2.2_R28) The rise of paper mills in all likelihood increases the (relative and absolute) amount of published research literature based on serious misconduct.
- > 2.2_O37) I believe some cases of serious research misconduct have been deliberately framed by journalists to attract public attention, creating a false impression that such cases are more common now than in the past. This has led to an unprecedented level of public visibility on the issue. Sites like Retraction Watch contribute to this perception...

Q. Effectiveness of official channels for investigating serious research misconduct

The strongest polarization of subgroups was in response to item 8.

Table 21. Item 8: "Official channels for reporting misconduct seldom work efficiently (select one option)"

Subitem	Percentage endorsing			
	1_S	2_R	3_O	All
Moderately/strongly agree	94.4	22.7	61.9	65.8

Throughout the free text responses, there was strong evidence of cynicism and lack of trust among some sleuths, relating to the previous point that they saw frequent instances where credible reports of fraud were not dealt with. The comments included many personal experiences that influenced these ratings:

- > 1.8_S10) I have reported on cases of academic misconduct without publishing in parallel on social media. Years later the universities had no idea that I had written to them, although they had acknowledged the receipt of the information...
- > 1.8_S9) It is hard to even find official channels in the first place. This could mean a complainant has had to contact several people about an issue before finding the right place/support, which means extra unnecessary exposure in situations that are often already complex and delicate.
- > 1.8_S5) The frustration is that in cases where the fabrication is so evident that a non-expert can see it in seconds, it still (sometimes) takes years to conduct an investigation, and, the correction of science is postponed (or just does not happen) even though it could and should happen quickly independently of establishing individual responsibilities (longer delays for that are quite understandable). One suggestion would be for institutions to transparently and publicly ask publishers for the publications of expressions of concern as soon as enough evidence that there are serious doubts about the reliability of the data have been provided.

In contrast, RI officers described systems working well. This of course, could be selection bias, as those who are motivated to take part in the survey are likely to be those with positive motivations to make the system work.

- > 1.8_R14) ...I have the impression our work is well-done, well-considered and the judgements well-founded. The slowness is mostly attributable to the fact that our investigation committee consists of about

ten full professors who meet physically and it is not considered feasible to have a meeting more often than monthly. When there's 350 pieces of evidence, this month between two meetings is well-needed.

- > 1.8_R22) ... I think the issue here is a matter of perception because sanctions/penalties are not broadly advertised so most people think that nothing happened. It would be great to consider how we can report on outcomes in a way that protects confidentiality of disciplinary processes...

For this item, as for many others, it was noted that there may be wide variation from place to place:

- > 1.8_R13) It is very hard to give a general answer here. Some countries and institutions have, in my view, accessible and relatively easy-to-use channels, others don't....

Further evidence of polarization on this topic came from responses to option (b) in item 6.

Table 22. Item 6b: "What is the impact of these factors in encouraging researchers to commit serious research misconduct? Please rate from 1 (little impact) to 5 (large impact)."

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
b) Low probability of being punished if detected and/or reported	77.8	40.9	52.4	60.8

One sleuth described it like this:

- > 2.6_S10) From my own experiences - the low probability of being detected and punished is an issue on multiple levels at universities, not only management, administration or integrity staff. It also concerns departments and even smaller teams. People in a department become aware, at some point, of some questionable practices or even misconduct committed by colleagues.... Even those who work directly with that person are unlikely to condemn their actions - even if they contribute to projects and knowingly participate in questionable practices contributing to potential misconduct...

R. Institutions handling investigations of their own researchers

Sleuths were strongly opposed to a self-regulation model. Many RI officers did not think the model of institutions handling their own cases was ideal, but noted two advantages: in principle it would allow institutions to learn by experience, and it also might make it easier to find investigators with knowledge of the subject area. Some respondents thought it might work best with support from an advisory body (such as is seen in the UK with the Committee on Research Integrity), which could give guidance to help ensure consistent best practice across institutions, and could monitor institutional statistics on this.

This was explicitly addressed in item 9, which asked for ratings about different options. (This item was poorly worded and some respondents thought they were required to rank order the options, which may have distorted results).

Table 23. Item 9a and 9c: "In an ideal world where resources are not an issue, which is the most suitable model/system for addressing serious research misconduct? Rate these options where 1 = least preferred and 5 = most preferred".

Percentage endorsing (4-5)

Subitem	1_S	2_R	3_O	All
a) Self-regulation approach, where academic institutions are responsible for conducting investigations and determining sanctions when one of their staff is accused	27.8	63.6	42.9	41.8
c) National guidelines approach; for example, UK's Committee for Research Integrity, which does not have a regulatory role, but aims to ensure all institutions work to a common framework	36.1	77.3	57.1	53.2

This respondent discussed pros and cons and explicitly noted how self-governance can allow institutions to learn lessons from cases they process:

- > 2.9_O27) ... I prefer self-regulation (with external quality control) because that is much better for learning and improving with a view to prevent future cases and also comes with less legalistic interference. Maybe we should work towards superinstitutional self-regulation by a disciplinary college, like we have for doctors and nurses.

One respondent felt the combination of institutional investigations with external guidelines would be best:

- > 2.9_O28) Only institutions can handle this efficiently, but having uniform external standards to guide institutions and possibly provide uniform sanctions would be an asset.

As noted above (Table 12), all groups noted that institutional conflicts of interest could be a hindrance in responding to serious research misconduct, so this was noted as a point of agreement. Nevertheless, there was a large gap in rates of endorsement: 91.7% for sleuths vs 52.4% for RI officers.

In free text comments, one RI officer reflected on the high rate of endorsement of COI concerns in Round 1, and noted that, even if COI was not a factor, the widespread belief that it affected decisions would be detrimental in addressing problems:

- > 2.5_R15) It is remarkable how high "Conflict of interest for institutions investigating their own researchers" was rated among Sleuths and overall.... Having been on the institutional research integrity office side, I have not seen that explicitly in my institution. More than a real problem, I guess the biggest impact of this self-regulation is that candidate-reporters have no trust in a hypothetical investigation and do not report after all. This is suggested also by what happened in Sweden: When Sweden established a national governmental board for research integrity investigations instead of regional offices they were overrun by an unpredictably high number of cases. Since 2020, they published 128 decisions so far.
<https://npof.se/en/decisions/>

S. A body similar to the UK Health and Safety Executive could be a useful model for research integrity investigations

This was one option given in item 9, which asked about the most suitable model/system for addressing serious research misconduct. The Health and Safety Executive differs from other options in that it could bring criminal prosecutions for serious research misconduct. It did not receive support overall, with RI officers being particularly opposed.

Table 24. Item 9g: "In an ideal world where resources are not an issue, which is the most suitable model/system for addressing serious research misconduct? Rate these options where 1 = least preferred and 5 = most preferred."

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
g) A body similar to Health and Safety Executive that could bring criminal prosecutions for serious research misconduct.	52.8	27.3	57.1	46.8

Comments on this item included:

1.9_S5) I would partly disagree with the comment on concentrating regulatory power in one location, if it is arguing for no centralised systems or sanctions - we have ISO standards and Health & Safety regulations for very good reasons, and certainly in the case of Health & Safety there is central body sanctioning power which has had positive impact (<https://theconversation.com/the-uks-health-and-safety-at-work-act-is-50-heres-how-its-changed-our-lives-235794>). There is a risk to assuming institutions acting as self-regulators are going to do a better job than examples set by established systems, and I could see this extending to misconduct regulations. A model similar to the current ISO accreditation methods (institutions are empowered in their quality process design, but there is still central oversight to specified standards) may be an option.

1.9_R17) ... I guess some specific harmful action could have criminal liability, like sharing sensitive information with malign foreign agents, harming the health or well-being of research subjects, serious mismanagement of public funds (e.g. corruption), fabrication of results with direct impact to safety, health and well-being of persons etc. But in case of plagiarism, it would be really difficult to imagine a case where criminal liability would be justified and necessary. I don't think that the concept of "research misconduct" fits well with criminal liability, even if we are focusing on the most serious forms, since the concept is (intentionally) broad and covers all sorts of unwanted practices. For conceptual clarity it would be better if a distinct concept or term was used for practices or deeds with criminal consequences – perhaps "fraud" or something similar.... "

1.9_R21) ... With regards to criminal prosecutions: this will be a lengthy process which won't give us the agility that these investigations require to fix matters of public record and address issues or research culture. It will be costly with little return on investment. Instead, could we explore models which already exist for nurses and medical staff where 'fitness to practice' can be removed or suspended.

T. Resources for research integrity investigations

The responses of sleuths tended to downplay the need for resources to deal with cases of serious research misconduct, whereas these were a major factor for those who conducted investigations, as evidenced by responses to Item 5.

Table 25. Item 5c: "Various factors may hinder academic institutions' response to serious research misconduct. Please rate the following from 1 (not much of a hindrance) to 5 (substantial hindrance)."

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
c) Lack of resources	38.9	76.2	47.6	51.3

U. Training of researchers

There was sharp division between sleuths and RI officers in responses to item 12, concerning attitudes towards training as a way to improve research integrity.

Table 26. Item 12b. "Given that we have finite resources, which solutions to serious research misconduct should be prioritised in funding?"

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
b) Research ethics training for all researchers	19.4	72.7	33.3	38.0

N.B. The failure to distinguish "ethics" and "integrity" training in the wording of this item was picked up by one respondent.

Some Sleuths noted that most of those committing serious breaches of research integrity knew perfectly well that it was wrong, but did it nevertheless.

For RI officers, training serves a useful function: it is a way for an institution to make explicit its commitment to research integrity, and is something where researchers can be monitored for compliance. This means that those who commit research misconduct cannot argue that they were unaware their behaviour was problematic.

Also relevant to this topic was response A of Item 15, which asked about options for institutions when serious research misconduct is confirmed.

Table 27. Item 15a: "Which of these practices should be options for institutions when serious research misconduct is confirmed. Please give your rating from 1 = strongly disagree to 5 = strongly agree".

Subitem	Percentage endorsing (4-5)			
	1_S	2_R	3_O	All
a) Educational retraining for researchers who have been found to commit serious research misconduct	33.3	86.4	52.4	53.2

V. Attitudes to those found to have committed serious research misconduct

Differing attitudes were not specific to individual items, but were evident in responses throughout the survey. The sleuths adopted a judgmental approach to those found guilty of serious research misconduct. Some noted that while factors such as pressure to publish might push some researchers into bad practices, in many cases, fraud was committed by people who were successful and had no need to conduct misconduct. It was also pointed out that pressure to publish affects everyone, but only some people respond with unethical behaviour.

- > 2.15_S6) It is too late to retrain people who publish fraud, they know what they are doing. Retraining could happen if you were caught faking data in an undergraduate report. One strike and you're out for published data. Either these people know what they're doing is wrong (so retraining will never work), or they are utterly stupid and unsuitable to be researchers in the first place.
- > 2.15_S8) ... Usually the approach is to protect senior individuals who commit serious misconduct when it is clear they know they have done wrong and to sack junior individuals who might have less insight.
- > 2.15_S4) I really don't think educational retraining would help. The misconduct does not happen because individuals are not well trained. They may be in fact very well trained in all aspects - research method, writing publications, research ethics, team management, science communication. It may be the excellent skills that enable them to spot opportunities for misconduct and ensuring that it's unlikely to be spotted/reported.

Responses from RI officers were more sympathetic. There were several comments noting the possibility of researchers being the subject of false accusations (several people commented "innocent until proven guilty"), or being swept up in the slipstream if a colleague or collaborator had engaged in misconduct. In general it was noted that when suspected misconduct was reported, this was highly stressful for the person involved, and it was important to offer support. For RI officers a popular option was re-education for those found to have committed misconduct - an option strongly rejected by the sleuths, who felt that such people had no role in research.

- > 2.13_R12) The principle of innocent until proved guilty should be embedded in processes.
- > 2.15_R18) Consequences and reporting/justice will depend on country, field of research and likely impact. There have been some discussion showing that the mental health strain of the worry about having been found to have carried out 'research misconduct' can be severe. I am concerned that there may have been (unreported) suicides or serious mental health damage to those accused of research misconduct that would not constitute criminality. Might it be good to carry out a study of the mental health impacts of allegations of misconduct and upheld allegations (where no harm was caused to the public by the misconduct and where no criminality occurred)?

Discussion

There were many points of agreement between sleuths and research integrity officers. Both groups endorsed the importance of tackling serious research misconduct, the need for support for individuals/organizations with expertise in detection, and the value of institutions and publishers/editors working together. They recognised academic incentives, the low probability of being detected/reported, and institutional conflict of interest as factors that encouraged serious research misconduct. The importance of having meaningful sanctions against serious research misconduct, and prompt action on removing erroneous work from the literature was recognised, as was the need to have protections for whistleblowers and collaborators who might be caught up in allegations of misconduct. There were two items that met the criterion for agreement, but where endorsement by sleuths was substantially stronger than endorsement by research integrity officers: the value of PubPeer as a source of information for identifying problematic publications, and the potential of an independent regulatory agency for dealing with allegations of serious research misconduct.

Against this background there were numerous points of disagreement. The frequency of serious research misconduct was estimated as serious by sleuths and less so by research integrity officers. This polarization mirrors the very disparate estimates of rates of serious research misconduct that are found in self-report surveys, versus the low rate of allegations to academic institutions (Bishop, 2025). Sleuths interpreted this to mean that only a small proportion of cases are detected, and indeed when they are detected the official channels for investigating allegations are wanting, whereas many research integrity officers felt that sleuths may have a distorted view of how much misconduct there was, influenced by a few cases that had a lot of media attention. Sleuths were opposed to self-regulation by institutions, noting that conflict of interest can make it difficult for them to take effective action (cf. Resnik, Hosseini, & Rasmussen, 2024; Schrag, Patrick, & Bik, 2025), whereas research integrity officers felt that self-regulation, supported by some kind of national body to specify guidelines and maintain some uniformity in procedures, was better than other options, which would have their own drawbacks. Another point of disagreement was whether training in research integrity was effective: sleuths expressed cynicism regarding this approach, and tended to attribute misconduct to deficiencies in the individual or distorted incentives, rather than to lack of knowledge. This related to the general negativity shown by sleuths towards those committing serious research misconduct, compared to more emphasis on due process and duty of care to those accused expressed by research integrity officers.

These views no doubt reflect the personal experiences of respondents with cases reported for serious research misconduct. Sleuths may have had experience of being ignored or attacked by individuals whose research publications contain unambiguous evidence of plagiarism, data fabrication or falsification, and their sympathy is accordingly low. High-profile cases of individuals such as Paolo Macchiarini, whose fraud led to patient deaths, are particularly upsetting, especially when, as in that case, the whistleblowers were attacked by a committee that initially vindicated Macchiarini (Elliott, 2024). Experiences of institutions failing to investigate allegations of research misconduct in a timely and transparent fashion may contribute to a cynical attitude (e.g., Grey, Bolland, Gamble, & Avenell, 2019).

Those conducting RI investigations may have sympathy for cases where a researcher has been put under strong pressure to produce publishable results in order to continue with their career. In addition, they may have encountered individuals in a state of distress at the prospect of losing their reputation and possibly their employment - at a point in an investigation when the evidence for misconduct is not clearcut. The need to protect the

mental health of anyone under investigation is highlighted by a case such as Howard J. Eisen, who took his own life after his lab was investigated by NIH (Eisen, 2013), even though he had not been accused of wrongdoing.

Conclusions

Siloed sources of evidence

It was striking how, in free text responses to statements in the survey, the sleuths and RI officers tended to provide different kinds of evidence. For sleuths, as well as personal experiences, the evidence consisted mostly of known cases of fraud, or survey data, reported either in the academic literature or in sources such as Retraction Watch, or commentaries in Nature or Science. The RI officers drew attention to a grey literature on the topic of research integrity, in the form of reports and policy documents. Indeed, some of them expressed frustration at the fact that these works, often the result of detailed labour by teams of individuals, were apparently unknown to those outside the policy sphere. Even within the RI community, there could be duplication of effort, as different groups set about assembling guidelines to address the same problems.

A difference between these two types of evidence is in the level of abstraction. Sleuths tended to refer to case studies that mostly illustrated the inadequacies of current systems. In contrast, formal policy reports aimed to specify general principles and procedures.

One might draw a parallel with the development of legal frameworks. Ultimately, the law specifies a set of general rules, derived by identifying principles that capture general points across a series of cases. We seem to have arrived at a point where formal frameworks are emerging but have yet to be adequately tested against real-world examples, so one can establish what does and does not work. One reason why this has not happened so far may be because of the secrecy surrounding research misconduct investigations; we cannot learn what does and doesn't work if there is no transparency. Improved interaction between individuals from different countries and constituencies could help progress on this point.

Rebuilding trust

There is an urgent need to restore trust between the two main communities that featured in the survey: sleuths and RI officers. It is not helpful for sleuths to demonize all those conducting RI investigations as insincere and corrupt individuals with no interest in getting at the truth. Likewise, sleuths will be unhappy at being characterized them as people who use the shield of anonymity to make unfounded or trivial criticisms against innocent victims. Undoubtedly, individuals exist who conform to each of these stereotypes, but we will only make progress if we accept that there are people in both camps whose sincere goal is to work to improve academic integrity, for the benefit of all. Nevertheless, the survey indicated this would be a difficult task because the items showing polarization on Round 1 tended to be just as polarized on Round 2.

Comparison of other organizational structures with parallel problems

This issue was not explicitly dealt with in the survey, but comparisons with other cases of organizational response to misconduct might provide the opportunity for "lessons learned". Within many organizations, including universities, there has been much criticism of inadequate responses to credible reports of sexual harassment, culminating in the Me Too

movement. There have also been well-publicized failures to address abuse of vulnerable individuals in religious organizations. There are some general characteristics of social structures within organizations that appear to foster inaction, including hierarchical power structures, a lack of clear lines of responsibility, closer identification with colleagues than with outsiders/victims, and a reluctance to break ranks within the group. Carl Elliott's 2024 book about whistleblowers, *The Occasional Human Sacrifice*, documents several cases where serious research misconduct was not dealt with adequately, where such factors were in play.

One question is whether we can learn from social scientists who have investigated group dynamics in these situations to ensure that our regulatory processes are set up to counter natural tendencies to avoid taking difficult actions (Bazerman, 2025). As Redman (2023) noted, policies in this area have tended to disregard blind spots that arise from aspects of human thinking and acting, such as a tendency to filter out information that destabilizes existing power structures and belief systems.

Examples of things that work

The survey provided an opportunity to express views on all the problems of the field: bad experiences, models that do not work, communication failures. Nevertheless, many respondents noted that we should not generalize from some bad cases to the whole research integrity field, given that there is widespread variation across nations, disciplines, and regulatory models. To move ahead, we need on the one hand to analyze cases where things have gone wrong, to identify the underlying causes, so that these can be mitigated in the future. But we also need to look at instances where things have gone well. It would be useful to collate examples of good practice, and use these to develop models that could be extended more broadly.

References

- Abalkina, A., Aquarius, R., Bik, E., Bimler, D., Bishop, D., Byrne, J., Cabanac, G., Day, A., Labbé, C., & Wise, N. (2025). 'Stamp out paper mills'—Science sleuths on how to fight fake research. *Nature*, 637(8048), 1047–1050. <https://doi.org/10.1038/d41586-025-00212-1>
- Bazerman, M. (2025). *Inside an academic scandal*. MIT Press.
- Berrio, J. P., & Kalliokoski, O. (2024). *Fraudulent studies are undermining the reliability of systematic reviews – a study of the prevalence of problematic images in preclinical studies of depression* (p. 2024.02.13.580196). bioRxiv. <https://doi.org/10.1101/2024.02.13.580196>
- Bishop, D. V. M. (2025). *A non-anonymised review of annual statements on research misconduct from UK institutions in 2023-4, with a focus on research fraud*. Zenodo. <https://doi.org/10.5281/zenodo.15373147>
- Bouter, L. (2023). Research misconduct and questionable research practices form a continuum. *Accountability in Research*. <https://doi.org/10.1080/08989621.2023.2185141>
- Carlisle, J. B. (2021). False individual patient data and zombie randomised controlled trials submitted to Anaesthesia. *Anaesthesia*, 76(4), 472–479. <https://doi.org/10.1111/anae.15263>
- Caron, M. M., Lye, C. T., Bierer, B. E., & Barnes, M. (2024). The PubPeer conundrum: Administrative challenges in research misconduct proceedings. *Accountability in Research*, 1–19. <https://doi.org/10.1080/08989621.2024.2390007>
- COPE & STM. (2022). Paper mills: Research report from COPE & STM. Committee on Publication Ethics and STM. <https://doi.org/10.24318/jtbG8IHL>
- Couzin-Frankel, J. (2015). PubPeer co-founder reveals identity—And new plans. *Science*, 349(6252), 1036. <https://doi.org/DOI: 10.1126/science.349.6252.1036>
- Elliott, C. (2024). *The occasional human sacrifice*. Norton.
- Gopalakrishna, G., Riet, G. ter, Vink, G., Stoop, I., Wicherts, J. M., & Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. *PLOS One*, 17(2), e0263023. <https://doi.org/10.1371/journal.pone.0263023>
- Grey, A., Bolland, M., Gamble, G., & Avenell, A. (2019). Quality of reports of investigations of research integrity by academic institutions. *Research Integrity and Peer Review*, 4(1), 3. <https://doi.org/10.1186/s41073-019-0062-x>

Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32(4), 1008–1015.
<https://doi.org/10.1046/j.1365-2648.2000.01567.x>

National Health and Medical Research Council. (2023). *International Research Integrity Policy Scan Final Report Compilation of information about research integrity arrangements outside Australia* [Purchase Order: PD003672]. [KPMG.com.au](https://www.kpmg.com/au)

Redman, B. (2023). *Reconstructing Research Integrity: Beyond Denial*. Springer.

Resnik, D. B., Hosseini, M., & Rasmussen, L. (2024). Responding to research misconduct allegations brought against top university officials. *Accountability in Research*, (online version) 1–6. <https://doi.org/10.1080/08989621.2024.2321179>

Sattler, S., Graeff, Peter, & Willen, S. (2013). Explaining the decision to plagiarize: An empirical test of the interplay between rationality, norms, and opportunity. *Deviant Behavior*, 34(6), 444–463. <https://doi.org/10.1080/01639625.2012.735909>

Schrag, M., Patrick, K., & Bik, E. (2025). Academic research integrity investigations must be independent, fair, and timely. *The Journal of Law, Medicine & Ethics: A Journal of the American Society of Law, Medicine & Ethics*, 1–4. <https://doi.org/10.1017/jme.2025.38>

Acknowledgments

This work did not receive any external funding. It was made possible by the generosity of moderators, the advisory board and panel members, who freely gave their time to make this project possible. The conference that was associated with this survey was funded by a research grant from St John's College, Oxford, with supplementary funding from Oxford University Press.

Declaration of interest statement

The author reports there are no competing interests to declare.

Data availability statement

The individual quantitative data used to create Supplement 6 are available in Supplement 9.