



*Promoting Integrity as an Integral Dimension
of Excellence in Research*

Printeger individual case study report

Deliverable Number	D 3.6
Work Package	WP III, Task T 3.3
Type	Report
Version	1
Number of Pages	15
Due Date of Deliverable	1 May '17
Actual Submission Date	July '18
Dissemination Level	Final, review by Serge Gutwirth

| Author

Willem Halffman

Case research by: Serge Horbach (Radboud University), Mari Rose Kennedy (University of Bristol), Sverre-Erik Mamelund (Oslo and Akershus University College), Marijke van Buggenhout, Jenneke Christiaens, Serge Gutwirth (Vrije Universiteit Brussel), Aive Pevkur, Kristi Lõuk, Mari-Liisa Parder, Marten Juurik, Katrin Velbaum, Heidi Meriste, Marie Soone, Kadri Simm, Margit Sutrop (University of Tartu), Ilaria Ampolini (University of Trento)



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 665926.

Table of Contents

1	Introduction	2
2	Design, methods and data	4
2.1	Research protocol	4
2.2	Unit of analysis	4
2.3	Methods and data gathering	5
2.4	Case selection	5
2.5	Data	6
3	Results	7
3.1	The multiple causation of misconduct	7
3.1.1	Performance pressure	7
3.1.2	Lack of social control	7
3.1.3	Status	8
3.2	Detection: who signals misconduct?	8
3.3	The articulation of misconduct	8
3.3.1	What is the issue?	8
3.3.2	Is it an issue? Normative ambivalence	9
3.3.3	Misconduct and conflict	9
3.4	Research organisations' initial response	10
3.4.1	Containment and backfire risk	10
3.4.2	The failure of ad-hoc solutions	10
3.5	Other institutional responses	11
3.5.1	The ambivalence of the media	11
3.5.2	The aftermath: regulatory responses	11
3.5.3	National differences	12
4	Conclusions and recommendations	12
	References	15



1 Introduction

In media and policy attention for research integrity, highly visible cases of misconduct have played an important role since decades (e.g. Broad & Wade, 1982). Accused or contested researchers and their stories form important reference points in the debate about what constitutes typical examples, causes and consequences of misconduct. Diederik Stapel (Abma, 2013; Stapel, 2016; Stroebe, Postmes, & Spears, 2012), Peter Nijkamp (Horbach & Halffman, 2017a), Jan Hendrik Schön, Hwang Woo-suk, Andrew Wakefield, or Paolo Macchiarini receive wide attention, while some even inspire fictional writing (Zwart, 2017). Even though such mediated cases may present a skewed representation of more common research misconduct, they still form important points of reference in the media (sic !), debate and policies around misconduct.

The accounts of mediated misconduct cases often focus on individualised aspects, such as culpability, the extent of misconduct, or individual psychology. For example, journalists attempt to discover how many of a culprit's publications are problematic, or which other researchers are involved (e.g. the work by *Retraction Watch*). In the Stapel case, the wrongdoer even briefly became a local media celebrity, testifying about how he had come to commit data fabrication, described in detail in an autobiography (Stapel, 2016). Along the same lines of individualisation, some researchers have attempted to identify the typical character traits of the scientific fraudster (Tijdink et al., 2016). The media's human-interest emphasis has certainly intensified this focus on individualised deviance causation.

However, it is also possible to see institutional processes at work in research misconduct cases, especially from a comparative perspective. Most accounts of misconduct focus on causation and damage, but they also offer a window onto how researchers get accused, how complaints are handled, or how (and whether) organisations adapt their policies in the aftermath. Such a perspective may offer a radically different view. For example, the book-length defence of two accused researchers in 'the Baltimore case' by historian of science Daniel Kevles denounces the institutional response for a misdirected, crime-fighting, scandal-seeking attitude that blinded investigators to the lack of substantial evidence of wrongdoing, which led to an eventual acquittal, but only a decade after the facts (Kevles, 2000). Meanwhile, outrage over 'the Baltimore case' had been used as one of the arguments to fuel stern regulatory intervention in US research (Guston, 1999). Clearly, cases of alleged misconduct can teach us more than just where individual misconduct comes from, but also how institutions respond – and sometimes miss the mark, falling short or overshooting adequate responses.

In addition, while media-reported cases mostly cover examples of spectacular wrongdoing, following the logic of scandal and revelation of clear 'black and white' transgressions (although with important exceptions), the cases that do not make the media also involve 'grey' cases of not-so-clear, not-so-spectacular questionable research practices (QRP). Systematic research into questionable practices shows that these are in fact much more common than outright fabrication, falsification, or plagiarism (Artino, Driessen, & Maggio, 2018; Fanelli, 2009; John, Loewenstein, & Prelec, 2012). Individual QRP misconduct cases may also become test cases for further articulation of norms, as debate among peers, media debate, or even disciplinary procedures become platforms for deliberation about what constitutes acceptable behaviour. An analysis of such less-exposed cases offers a perspective on the effects of mediated controversy around research misconduct.



For these reasons, the PRINTEGER project included a series of individual misconduct cases studies, two for each project partner with research capacity. This resulted in cases studies from six countries. As there were two partners from the Netherlands and since the work of the Leiden University partner focused on the operation of the publication system, the Leiden contribution focused on more detailed work on journals and the publication system. The result of this work is not included in this report, but is documented separately (deliverable D3.5 case studies of predatory publishing and retractions) (Reyes Elizondo, De Rijcke, & Van Leeuwen, 2017).

Identifying and getting access to individual cases of misconduct proved much harder than was anticipated on the basis of a preliminary scan for the Netherlands when the proposal was written. While in the Netherlands an annual report of anonymised cases reported through integrity committees is publicly available (<https://www.lowi.nl/nl/over-lowi/jaarverslag>), this turned out to be an exceptional practice. In other countries, misconduct cases are treated with far more confidentiality. In the UK or Belgium, for example, not even summarising statistics of reported misconduct cases are available. Hence we had to make some compromises over the kinds of sources we could work with and had to rely on publicised material only for some countries.

The Printeger project also intended to use the material gathered about the individual cases as illustrative material for the educational tool developed in work package five, the 'UPRIGHT' e-learning package (<https://upright.science.ru.nl/>). Eventually, only the case studies based on public sources could be used for these purposes, given the confidentiality conditions around most cases. Cases accounts that were cleared for public availability are included in the annex to this document.

The full objectives of the cases studies were therefore:

- "To gather illustrative material for use in educational tools for early career scientists (in WP5). Hence cases are not just about 'pointing out obvious wrong behaviour', but illustrate the dilemmas and tensions in research that lead to dubious practices, the tensions between different principles or guidelines, point out how misconduct cases can develop (e.g. the dynamic of 'scandal'), or the complications of whistle blowing.
- To look for patterns in the development of misconduct cases in terms of the personal dynamics of fraud, in light of the tensions arising from work conditions, such as pressures or research culture. (A specific focus in light of information gathered in the context of other work packages.)" (Halffman, 2016)

The cases studies specifically probed patterns with respect to:

- Origins and causes of misconduct, including organisational setting and institutional pressures;
- The official processing and institutional response to the allegations, along with the range of actors that got involved;
- How the case was interpreted and how it was framed, including in the media;
- Consequences and outcome of the case, not just for the individual, but especially on organisational and even institutional level.
-



This document describes the methods and research design, the case analysis and the main conclusions from the cases. Case descriptions suitable for public access and use in teaching are included in the Upright tool.

2 Design, methods and data

2.1 Research protocol

The design of the case studies was developed jointly with the project partners in several discussion rounds during project meetings. Issues of data access, comparability and ethics of reporting on misconduct cases were major issues. While the Netherlands has publicly available (but anonymous) overviews of reported cases, any public account of misconduct cases is considered extremely sensitive in the UK. In Belgium too, misconduct cases that go through disciplinary procedures are shielded from any kind of public access. We decided to work with what was locally acceptable and available.

While the basic research ethics issues (about informed consent for interviews, privacy, information removal) had been dealt with, further ethics issues appeared as research plans became more concrete. One such issue concerned our position as researchers if we should have discovered further misconduct in our case accounts. Some team members signalled that their country's legal requirements would oblige them to report such information, while other researchers with experience in studying deviance argues that the position of a research is not compatible with a policing role. We agreed to decide in such cases as they would occur, with a veto right on reporting by research team members. In the end, no such cases occurred, but the issue does show the complications of doing case research in this field.

In addition, from a methodological perspective, the team had to decide on how to make cases as comparable as possible under conditions of extremely difficult and sensitive data access, and on the other hand decide which parameters could be compared and contrasted between cases. While data collection had to be pragmatic, the team did have to reach agreement about what constituted a case, about where case descriptions should begin and end, and about how cases should be reported.

The eventual agreement over the design of the case studies was articulated in a research protocol. The research protocol provided the definition of cases, principles for case selection, data sources, processes to pay attention to and indications for the case narrative, as well as practical instructions about length and format (Halfman, 2016). While this protocol had considerable benefits for the comparability of the case narratives, in some cases the confidentiality surrounding misconduct made it very hard to gather even basic information. In fact, the considerable variation in how misconduct cases are reported and documented in the various national systems for dealing with misconduct is an interesting observation in itself (see below).

2.2 Unit of analysis

Cases were defined as individual cases of researcher misconduct (to repeat: with the exception of the Leiden cases about publishing, excluded here). A case was defined as a story of events building up to misconduct or alleged misconduct, the allegations, the response by colleagues and organisational environment, as well as resolution of the allegations and, where relevant, wider repercussions of the case. Wider repercussion could include, for example, the articulation of rules, improved policies or procedures, or changed agenda setting.



Extending the case into the origins of misconduct as well as further repercussions (beyond the procedural adjudication) proved difficult for cases that were not exposed in the media. Only in the Netherlands and Norway, and to a lesser extent Estonia, was it possible to develop less exposed cases through additional interviews. For Belgium in particular, with hardly any options for a case of misconduct reported in the media, the cases had to remain largely restricted to the part of the story documented in the case files, for which researchers acquired access under very strict conditions of confidentiality.

2.3 Methods and data gathering

As access to individual misconduct case information proved extremely difficult, a mix of methods had to be used, relying on the analysis of official (that is: academic) case files (Belgium); case files, secondary analysis of reports and media accounts, combined with interviews (Netherlands, Norway, Estonia), to mostly secondary analysis of already published materials in media or scientific papers and some limited interviews (UK, Italy). To foster maximum comparability of the cases, the research protocol specified particularly on what should be considered part of a case, but out of necessity left the methods to the research teams, suggesting only what kind of documents or interviewees could be consulted. Details can be found in the research protocol (Halffman, 2016).

2.4 Case selection

Case selection was also co-ordinated between the partners, although for the UK and Italy in particular, we were much limited (and admittedly biased) by what was known in the media. Two sets of criteria were articulated for case selection for comparative purposes: one set focusing on representative properties, one set focussing on contrasting properties. For representative purposes, we tried to get a spread of gender, disciplines, types of misconduct, academic/non-academic institutes, and countries (as far as represented in the project). Non-academic research was even harder to access and we had to abandon that criterion. For comparative purposes, we aimed to contrast high and low status researchers (high being professor, senior researcher and low being junior researcher, PhD or similar), high and low levels of public attention (media), and clear versus contested forms of misconduct.



2.5 Data

Eventually, thirteen cases were analysed. Due to extremely limited data, the Belgian partner offered one extra case.

partner	code	Comparative			Representative			
		status	Visible	Contested	field	gender	instit	form
RU	NL1	high	high	High	Economics	male	univ	Tex recycling
RU	NL2.	low	low	Low	Social Science	female	univ	Plagiarism
VUB	BE1	high	some media (name)	Low	criminology	male	univ	Plagiarism
VUB	BE2	high	low, no media	Moderate	medicine	male	univ	Dual publication
VUB	BE3	high	low, no media	Moderate	Social science	?	univ	Fabrication
CEUT	Est1	Low phd stud	Low	Grey	Education	female	univ	Accused plagiarism
CEUT	Est2	High (senior)	High	?	Economics	male	univ	Text recycle, predatory publishing
HIOA	No1	high	high	High	Medicine and odontology	male	univ	Fabrication
HIOA	No2	low	low	Low	Teaching/pedagogical sciences	Not known	Private univ college	Plagiarism
UT	It1	high	High (media, Parliament)	Yes	Veterinary science, GMO	male	Univ	Data falsification
UT	It2	high	High (TV, newspapers)	Yes	Medical	male	Univ+ hospital	various
UBris	UK1	Clinician & research fellow – medium /high	High	Clear	Medical/biology	male	Aca-demic	Conflict of interest, question. ethics, falsification
UBris	UK2	Professor, senior	high	Moderate	Health sciences-	male	Aca-demic + indus.	Sloppy science , commercial

Table 1 Overview of cases



3 Results

Several processes can be elucidated through comparison of the cases, although not as extensively as we had hoped prior to the research. What follows is a list of patterns that can be observed in the cases, ordered along a set of common themes. Among the cases that could be studied by us in most detail, were the case studies for Norway and The Netherlands, a comparison that will result in an academic publication (Horbach, Breit, & Mamelund, 2018 forthcoming). In the text below, cases are referred to using the code mentioned in Table 1.

3.1 The multiple causation of misconduct

3.1.1 Performance pressure

Individual cases stories obviously do not offer the best source for hard evidence on the systematic causes of research misconduct. For example, processes that appear in several of the cases are work pressure and compelling research evaluation requiring 'output'. Researchers may need publications to graduate, advance careers (their own or the career of students and team members), or maintain support for research groups (e.g. when finances depend on 'output'). Evidence that such evaluation practices can 'over-shoot' and cause systematic undesirable research behaviour is better gathered through other research methods (Fanelli, 2010; Hicks, Wouters, Waltman, de Rijcke, & Rafols, 2015), but nevertheless also appears as a theme here.

Pressure to produce is flagged in several of our cases. In Est1, this involves pressure to finish a PhD project in time. In UK2, it involved a race for priority in discovery. Estonian investigations led to evidence that performance pressure had lead researchers to publish in predatory journals to achieve 'international' publication output, and responses from researchers that suggest 'gaming' behaviour became normalised (Est2).

However, if anything, the cases also show that causation is complicated. For example, researchers still engage in text recycling even when hard pressure to perform is absent, with no hard financial reward for high publication rates, or senior researchers that no longer need high publication rates (e.g. NL1). There are strong indications that some questionable research practices are maintained by the research culture of a particular field, as much as of 'hard' organisational performance pressures. For example, a more detailed analysis of text recycling triggered by case NL1 showed that rates of text recycling vary considerably between research fields and that some journal editors seem to have little objection against publication of recycled material (Horbach & Halfman, 2017a). Other research has found that cynicism over procedural injustice, such as in unreasonable assessment of research quality, can reinforce a culture that justifies misconduct or questionable practices (Clair, 2015). Without downplaying the importance of institutional conditions such as evaluation pressures, the causes of misconduct are clearly complex.

3.1.2 Lack of social control

In several cases, actors signal a lack of social control as the cause of misconduct. Examples are co-authors that did not read drafts of a paper (No1); are not aware of research details and did not check the data on which papers were based (UK1); or did not even have access to the data (UK2); or co-authors that did not notice data fabrication or missing ethics approvals (Be2). Ritual authorship can implicate researchers in misconduct, rather than providing the critical check on work that could be provided by co-authors. Some of the cases researched could have been prevented if cooperating researchers had kept a closer eye on each other, including data and drafts.



However, the mere fact that multiple researchers are involved or look at each others' work is no guarantee: cases also offer examples of 'group think' (It1), or research cultures that have come to accept allegedly questionable practices such as text recycling as normal (NL1). Social control may therefore be more than just a matter of oversight by informed insiders, but may also require contributions from critical outsiders.

3.1.3 Status

Several of the cases involved researchers in prominent positions (see Table 1). In fact, several of our cases involved research 'super-stars' (e.g. NL1, No1, It1) of high prominence in research and in some cases also beyond, as public figures. While over-representation is most probably caused by our forced reliance on high-profile media cases (which may be extra-interested in cases involving prominent researchers), also cases that came out of case files (e.g. the Belgian cases) show that misconduct is not just a matter of junior researchers unaware of constraints or insufficiently socialised in the practices of their research field. This raises questions in light of policy measures that attempt to educate young researchers in research integrity, while some of the problems may be caused by behaviour of their role models.

3.2 Detection: who signals misconduct?

While several of our cases were brought to the attention of integrity boards or committees by peers and colleagues, whistle blowers also took cases to journalists (e.g. NL1, It1, It2), typically in an attempt to raise attention for issues that were not taken up by internal procedures. In such cases, status complicates detection, as high status researchers get exceptional treatment such as fast-tracked papers (No1), or through the career dangers for junior researchers when challenging high-status seniors.

In addition to these well-known channels, journals also appear in some of our cases. For example, journals may be informed of misconduct cases through calls for retractions, which in turn may come from research peers, main authors themselves, co-authors, or research organisations (as reported regularly by *Retraction Watch*). Reporting misconduct may present a burden on editorial capacity, especially if editors want to make sure reported cases are followed up (Martin, 2013). However, journals may also have a particular interest in research integrity cases, for example in the protection of exclusive control over copyrighted material to protect their business model. This was at stake in one of the Belgian cases, revolving around an accusation of 'redundant' publication by a publisher (Be1).

3.3 The articulation of misconduct

3.3.1 What is the issue?

As misconduct allegations are made, a process of articulation starts: attempts to identify if and what misconduct is at stake, or how misconduct is entangled with a range of 'other' issues. Whereas some cases are relatively simple, singular forms of misconduct (e.g. straight-forward plagiarism, such as NL2), most of the cases we researched presented more complexity. Complex cases involve diverse allegations, such as concerns over plagiarism *and* concerns over the quality of work (Est1, NI1). With questions about quality sometimes comes disagreement over proper methods (UK2), tensions between schools of research, especially in social sciences where "proof" is often considered problematic, or what constitutes good research in the first place. Furthermore, concerns over integrity may lead to discoveries over research ethics, falsified CVs, even accusation of criminal behaviour (extortion, It1). Research misconduct may also be mixed in with work floor conflicts (either



as cause or consequence), or be presented as integrity conflicts in order to be acceptable to an integrity board, perceived as a possible forum to settle a conflict (No2). In such cases, actors may disagree whether an allegation is an integrity issue in the first place.

From a procedural point of view, splitting a case into component parts may present a way to make it actionable, containable: separate the integrity issue from the work floor conflict, from the research ethics, or other aspects. This makes the contentious issue suitable for specific arenas to resolve the issue: integrity boards, managerial intervention, the research ethics committee, or possible courts. Such distribution makes cases procedurally manageable, including the provisions that protect the accused by offering fairness or proportional penalties. At the same time, splitting up a case may make it harder to get to the heart of a conflict, to provide conflict mediation, or to deploy effective organisational or policy interventions to prevent future problems.

3.3.2 Is it an issue? Normative ambivalence

Several cases involved contention over what constitutes misconduct and of alleged behaviour actually constitutes misconduct and how this may vary between research fields or types of professional work. For example, in Est1, the definition of plagiarism was at stake, as there was dispute of when precisely copied text actually constitutes plagiarism, or what forms of referencing to sources are required or appropriate. In No2, differences conceptions between pedagogical professionals about referencing original material were at stake, while in It1, there was debate between clinicians and researchers about acceptable methodologies and evidence standards. The ambivalence extends to debates about the very status of ethical principles, for example in whether all that is not expressly forbidden is allowed or what consequences should be of unethical behaviour that is not breaking a rule (No2), whether and when negligence becomes a breach of integrity, such as when wrongdoing cannot be proven due to inadequate data storage (Be3).

Normative ambivalence seems to be a rule rather than the exception in our cases. Perhaps cases with less ambivalence get resolved at lower organisational levels (before they become visible in procedures or media), or ambivalence is created as a defence strategy by those accused, producing a bias resulting from how we looked for cases. Nevertheless, while more precise articulation of norms may reduce some of the normative ambivalence, the complexity of research work and diversity of research cultures is not likely to be completely resolvable by more detailed rules, underlining the importance of professional judgement in deciding on integrity breaches.

3.3.3 Misconduct and conflict

Conflict was a recurrent theme in several of our cases. Conflicts may be about opposing (economic) interests, such as between a journal's copyright and authors' interest to share knowledge (Be2). In some cases, the issue was framed as a misconduct case possibly as an outlet for work floor conflict (e.g. No2). Conflict between whistle blowers and accused researchers may be about personal issues, but also typically involves a degree of professional indignation: the whistle blower feels professional values and standards need to be defended. Research standards (NI2) or fairness of academic rewards (NI1) are perceived to be under threat in these cases.

Moral indignation clearly fans some of the conflicts, in some cases also involving journalists appealing to outrage among their readers (e.g. NI1, It1). Moral outrage particularly seems to play a role in protracted conflicts that fail to be resolved by investigative committees. While understandable and perhaps justified as a way to confirm professional standards or shared values, such outrage also



contains the risk of overreaction and scapegoating. In some of our cases, this has arguably led to overreactions, with attempts to exile alleged wrongdoers from their professional careers, even for minor wrongdoings or for breaking very implicit or ambivalent rules (No1). These cases underline the importance of accommodating conflict resolution and preventing scapegoating in integrity procedures.

3.4 Research organisations' initial response

3.4.1 Containment and backfire risk

The reflex of research organisations to immediately contain and preferably minimise misconduct cases is remarkable. Fear of media exposure and reputational damage to the host organisation is a major factor in how misconduct cases are initially dealt with, especially if whistle blowers are involved. Fear of media exposure can go as far as cover-up reactions (Est1, UK2). Even scientific journals seem hesitant to recognise misconduct is at stake in articles (UK1), behaviour regularly confirmed in the reporting by *Retraction Watch*.

However, attempts to minimise or cover up misconduct may well backfire. If whistle blowers who took a risk by signalling problems feel they are not taken seriously, they may pursue the matter further, for example by involving the media. This may then lead to even larger media exposure, with more reputational damage for the research organisation involved (It1, UK2). Denial, minimising, or covering up misconduct cases therefore has a risk of backfiring if somebody feels the need to pursue such cases further.

Research organisations also display particular containment strategies. One is to reduce complex or systematic integrity breaches to an individual's behaviour (e.g. NI1). By reducing the problem to an individual 'black sheep' or 'rotten apple', research organisation hope to limit the damage and 'solve' the problem by blaming individual researchers (Van Buggenhout & Christiaens, 2016). A similar containment strategy is to compartmentalise the issue, such as by separating 'research integrity breaches' from other wrongdoings, such as criminal or personal problems, for which the organisation can then deny responsibility (It1). While such containment strategies may limit reputational damage for the organisation, they also fail to address potential systemic causes or more widespread questionable practices and may obstruct policies that address more fundamental problems and prevent further cases.

3.4.2 The failure of ad-hoc solutions

A typical reaction of a research organisation facing unfamiliar research misconduct without appropriate procedures is to set up ad hoc investigative committees, usually consisting of in-house senior researchers. Such committees are then tasked with investigating the extent of misconduct, establishing whether norms have been broken, determining culpability and recommending potential disciplinary measures (NI1, No2, UK1, No1, It2).

Generally, this does not go well. Ad hoc committees tend to lack experience in misconduct cases and hence run a high risk of making mistakes of a procedural or legal nature, such as disregarding fair procedural guarantees for the accused researcher. They may also find they lack effective means to sanction or otherwise intervene in a case. Especially prominent researchers often fight back, seeking legal support to challenge the actions of ad hoc committees and the organisations that installed them – and with some success (It2, NL1, No1, UK1). Challenges sometimes lead to sequel ad hoc



committees, which suffer from the same problems. Ad hoc committees therefore tend to fail in providing closure, sanction, and fair treatment of accused researchers.

The result of messy ad hoc procedures may be an awareness that more formal procedures with adequate safeguards are required, but several misconduct cases would have been less traumatic had fair and adequate procedures been in place before they came to light.

3.5 Other institutional responses

3.5.1 The ambivalence of the media

Media exposure offers a most ambivalent channel to address research misconduct. On the one hand, the media may offer a back-up channel if misconduct issues have been neglected or if research organisations fail to take appropriate action. In several of the cases we studied, the media investigated cases (or provided a forum to investigate cases) that had been left undetected. In this sense, media exposure can be a channel for whistle-blowers to signal trouble after internal procedures failed to resolve misconduct, or are perceived as such (e.g. NL1, It1, It2). In UK1, the media criticised research organisations for not addressing misconduct, including journals for not retracting problematic papers. In this sense, media attention can act as a powerful check on researchers, research organisations, research governing institutions, and even research journals.

However, media exposure can also work in less productive ways. For example, the lens of the media may lead to over-simplification of misconduct, for example by overstating cases in a search for a strong media story, or through a focus on individual human interest. Making misconduct a matter of individual psychology may reinforce containment of misconduct as isolated, individual effects. Nevertheless, stories of misconduct in the media do also connect misconduct to more structural features, such as pressure from funders or performance pressure (Ampollini, Bucchi, & Saracino, 2016).

More worryingly, in UK1, the media also offered a forum for the culprit, with a channel to announce 'discoveries' parallel to peer review and in stronger terms than would be acceptable in a professional forum, but also as a means to fight back against proven allegations. In NL1, the media became a platform to denounce an alleged culprit, who was vilified but then later found innocent in court. This is not just a matter for reputational damage for the researcher accused of misconduct, but also the whistle blower can become a target (UK2), while some media's search for drama may exaggerate the credibility of discredited research. While the media have an important role to play in the democratic oversight of research, there therefore are also good reasons to be wary about media exposure of misconduct cases.

3.5.2 The aftermath: regulatory responses

Several of our cases led to regulatory interventions, especially some of the big cases that made the press. Through the eruption of misconduct cases and the debate they give rise to, norms may get clarified, either as cases become precedents, or as research organisations issue more precise rules and procedures (NI1, No1). In response to visible misconduct cases, regulations and guidelines become more articulate and elaborate; also as research organisations and research governing boards display their commitment to restoring confidence in research and maintain professional standards. (In some cases, such improvements were promised but eventually not implemented, Est1.) Specification of rules and procedures is a way to show that breaches are taken seriously, that future breaches will be dealt with sternly, and that organisations are in control.



Some case also led to the introduction of courses or other policy instruments to improve research integrity or prevent misconduct, but the dominant reaction is to go for regulatory initiatives: to articulate rules and implement procedures to maintain them.

3.5.3 National differences

Our cases show that there are considerable national differences in how research conduct is dealt with and in the institutional framework addressing research misconduct and integrity. In some countries, procedures and research integrity committees are virtually absent, with research organisations relying on the traditional approach of senior scientists' committees or hierarchy in research organisations to deal with misconduct (e.g. Italy). In other countries, research organisations have research integrity procedures and committees on the level of research organisations, but not external or national appeal boards (e.g. UK). Other countries combine integrity committees in research organisations with a national committee that acts as an appeal board and that collects information about cases (e.g. The Netherlands), or as a provider of a non-binding 'second opinion' to the decision-takers in the research organisation (e.g. Flanders/Belgium). Yet other countries have a stronger role for the state and the courts, with specifications of research misconduct in criminal law (Norway).

4 Conclusions and recommendations

This report only presents part of the puzzle for the Printeger project. While there are strong qualitative observations to be made from the case studies, the observations need to be combined with other parts of the research (survey, focus groups, legal and normative analysis).

Nevertheless, the information gathered in this part of the project leads to conclusions that also support some specific recommendations:

Causation: multiple

Beyond individual origins of misconduct, organisational processes and science policies play a role, but these cannot be reduced to single factors. While there are signs that excessive performance pressure may encourage researchers to cut corners, also a lack of social control and mutual oversight, as well as processes of group think or mutual confirmation can play a role, even on a research field level. Research integrity is also a matter of a shared culture, while inversely a culture of cynicism (Clair, 2015) may induce questionable practices, for example as behaviour 'gaming the system' becomes normalised because evaluation and promotion procedures only take output quantities and numbers into consideration. In any case, an integrity policy that addresses the generation of research misconduct as well as dealing with transgressions will require a wider perspective than just single causes or policing action. This includes particular attention for status inequalities, role models in research and the expectations and thresholds generated by science policies .

Recommendation: Integrity policy has to take into account the complex generation of misconduct. Prevention of misconduct through one factor alone is unlikely to be effective, just as research integrity should not be considered as just a problem of socialisation for young researchers. The adverse effects of the evolution of science policies towards competitive knowledge production systems should at least be taken serious as well.

Detection: clarify journals' role



The obvious ways to signal cases of misconduct run through internal procedures (along the hierarchical line, or to research integrity boards), or external whistle blowers. The importance of proper procedural guarantees for such cases has been a main feature of integrity policy. However, the role of journals in integrity policy is more contentious. The responsibility of editors to signal integrity, the expectations of editorial effort, as well as the interest of journals themselves has received little attention and requires further clarification. The role of journals is analysed in more detail in other parts of the Printeger project (Reyes Elizondo et al., 2017).

The issue: separate integrity?

While procedural guarantees (including fairness to the accused) and conflict containment may require that research misconduct is separated from disputes over research quality, ethics, or personal conflicts, actual conflict resolution and prevention policies may require that cases are considered in their complexity, rather than reduced to ‘the proper box’.

Recommendation: While for the purpose of fair procedures research misconduct should be well demarcated and properly defined, research integrity *policy* cannot separate research integrity from closely related issues such as research evaluation, levels of competition, organisational culture, or transparency and the possibility of open intellectual debate.

Misconduct articulation and resolution

Misconduct cases involve more than just broken rules: they also involve normative ambivalence, disagreement, conflicts of interest, conflicting interpretations, or moral outrage. The resolution of misconduct allegations is therefore more than just establishing whether a rule has been transgressed, but also contains an occasion for articulating norms or principles, and a need to resolve or de-escalate conflict. From a policy perspective (in organisations or in research systems), misconduct may be connected to other processes, such as problems in organisational culture, career or research evaluation, or in the uneven access to resources.

Recommendation: While ruling over individual misconduct cases in order to establish culpability is a crucial element in research integrity policy, such policy also needs to pay attention to conflict resolution and systemic connections between integrity and other processes, such as research evaluation, resource distribution, or organisational culture.

Organisational response: denial is unwise

While the reflex reaction of research organisations to minimise misconduct cases is understandable in the face of possible reputational damage, the backfire effect of failed suppression may well be far worse. The more effective strategy is likely to be based on honest recognition of wrongdoings and a clear demonstration of a willingness to take action. Nevertheless, some form of containment of misconduct cases is required, particularly from the perspective of procedural guarantees of both fairness and effective intervention or to limit conflict, even if misconduct cases can give rise to more profound reflection on fundamental work conditions or research cultures further down the line.

Recommendation: From a long-term perspective, the reflex of research organisations to minimise or deny misconduct is not a wise strategy, even though some containment may be required.

Media

The role of the press in research misconduct is ambivalent. While on the one hand the media offer a back-up channel for individual cases that were inadequately dealt with, they also contain the risk of



fanning conflict, offering a podium for discredited research, or damaging reputations unjustly. The presence of research super stars among cases that achieve media notoriety may become reference points for integrity culture, but with an ambivalent message: some researchers have built successful research careers and got away with questionable practices, while the vilification of others reconfirms research community norms, but possibly at the expense of fair treatment.

Regulatory policy response

The dominant response to misconduct cases to articulate rules and implement procedures to enforce them (and enforce them more fairly) is understandable and may be a way to make ‘good use of a crisis’ to improve the normative framework dealing with research integrity. However, in terms of policy, such responses tend to reach for the ‘stick’ and neglect the ‘carrot’ and the ‘sermon’ among the broad categories of policy instruments (Bemelmans-Videc, Rist, & Vedung, 2010). This pattern was also signalled in our analysis of research integrity discourse: policy documents have developed more in the direction of maintaining rules or ‘policing’ research integrity, along with a growing discrepancy with the kinds of integrity policies preferred by researchers (Horbach & Halfman, 2017b). While strong regulatory action may show willingness to restore moral order, it may not be the only or even the best way to address the sources of research integrity problems.

Recommendation: Research organisations and governing boards could consider a wider pallet of policy instruments than just articulation of more specific rules and procedures to enforce them.

National differences and the need for procedures

Such national differences partly result from how the debate over research misconduct scandals has played out, but they also reflect deeper institutional differences in how polities operate. While there are good reasons to learn from experiences in other countries, it would make little sense to try and recommend or even enforce one model. This would disregard national institutional differences and ultimately even the freedom of a political society and/or and a scientific community to shape itself. Nevertheless, it seems clear that research integrity requires *some* form of institution that can articulate research integrity norms and principles, and can guarantee effective and fair procedures for dealing with cases. Ad hoc solutions (such as ad hoc investigative committees) have proven painfully inadequate for dealing with individual cases in the past. Several examples for procedures or committees are now available in a wide range of national contexts.

Recommendation: The absence of any kind of formal research integrity committee to deal with cases is no longer an option, although such committee can be organised in a wide variety of ways, as demonstrated in different national contexts.



References

- Abma, R. (2013). *De publicatiefabriek. Over de betekenis van de affaire Stapel*. Nijmegen: Uitgeverij Vantilt.
- Ampollini, I., Bucchi, M., & Saracino, B. (2016). *Report on media analysis*. Retrieved from <https://printeger.eu/wp-content/uploads/2017/10/D3.2.pdf>
- Artino, A. R. J., Driessen, E. W., & Maggio, L. A. (2018). Ethical Shades of Gray: Questionable Research Practices in Health Professions Education. *Bioarxiv*. doi:10.1101/256982
- Bemelmans-Vidéc, M.-L., Rist, R. C., & Vedung, E. (Eds.). (2010). *Carrots, Sticks & Sermons: Policy Instruments & Their Evaluation*. New Brunswick: Transaction.
- Broad, W., & Wade, N. (1982). *Betrayers of the Truth: Fraud and deceit in the halls of science*. New York: Simon & Shuster.
- Clair, J. A. (2015). Procedural Injustice in the System of Peer Review and Scientific Misconduct. *Academy of Management Learning & Education*, 14(2), 159-172.
- Fanelli, D. (2009). How Many Scientists Fabricate and Falsify Research? A Systematic Review and Meta-Analysis of Survey Data. *PLoS ONE*, 4(5), e5738.
- Fanelli, D. (2010). Do Pressures to Publish Increase Scientists' Bias? An Empirical Support from US States Data. *PLoS ONE*, 5(4), e10271. doi:10.1371/journal.pone.0010271
- Guston, D. (1999). *Between politics and science: assuring the productivity and integrity of research*. Cambridge: Cambridge University Press.
- Halfman, W. (2016). *Misconduct cases: research protocol*. Retrieved from Nijmegen: <http://printeger.eu/wp-content/uploads/2018/06/3.4-case-studies-protocol.pdf>
- Hicks, D., Wouters, P., Waltman, L., de Rijcke, S., & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, 520(7548), 429-431. doi:10.1038/520429a
- Horbach, S. P. J. M., Breit, E., & Mamelund, S.-E. (2018 forthcoming). Organisational responses to alleged scientific misconduct: Sensemaking, sensegiving and sensehiding. *Science and Public Policy*.
- Horbach, S. P. J. M., & Halfman, W. (2017a). The extent and causes of academic text recycling or 'self-plagiarism'. *Research Policy*. doi:10.1016/j.respol.2017.09.004
- Horbach, S. P. J. M., & Halfman, W. (2017b). Promoting Virtue or Punishing Fraud: Mapping Contrasts in the Language of 'Scientific Integrity'. *Science and Engineering Ethics*, 23(6), 1461-1485. doi:10.1007/s11948-016-9858-y
- John, L. K., Loewenstein, G., & Prelec, D. (2012). Measuring the Prevalence of Questionable Research Practices With Incentives for Truth Telling. *Psychological Science*, 23(5), 524-532. doi:10.1177/0956797611430953
- Kevles, D. J. (2000). *The Baltimore case: A trial of politics, science, and character*. New York: W.W. Norton.
- Martin, B. R. (2013). Whither research integrity? Plagiarism, self-plagiarism and coercive citation in an age of research assessment. *Research Policy*, 42(5), 1005-1014. doi:<http://dx.doi.org/10.1016/j.respol.2013.03.011>
- Reyes Elizondo, A., De Rijcke, S., & Van Leeuwen, T. (2017). *Handling publishing misconduct: tools used by publishing houses and editors*. Retrieved from Leiden:
- Stapel, D. (2016). *Ontsporing*. Amsterdam: Prometheus.
- Stroebe, W., Postmes, T., & Spears, R. (2012). Scientific Misconduct and the Myth of Self-Correction in Science. *Perspectives on Psychological Science*, 7(6), 670-688. doi:10.1177/1745691612460687
- Tijdkink, J. K., Bouter, L. M., Veldkamp, C. L. S., van de Ven, P. M., Wicherts, J. M., & Smulders, Y. M. (2016). Personality Traits Are Associated with Research Misbehavior in Dutch Scientists: A Cross-Sectional Study. *PLoS ONE*, 11(9), e0163251. doi:10.1371/journal.pone.0163251
- Van Buggenhout, M., & Christiaens, J. (2016). *Deviance in science: a criminological analysis*. Retrieved from Brussels: <http://printeger.eu/wp-content/uploads/2016/12/D2.5.pdf>
- Zwart, H. (2017). *Tales of Research Misconduct: A Lacanian Diagnostics of Integrity Challenges in Science Novels*. Springer Open.