Explanations of Research Misconduct, and How They Hang Together

Tamarinde Haven¹ · René van Woudenberg¹

Accepted: 12 January 2021 / Published online: 19 May 2021
© The Author(s) 2021

Abstract

In this paper, we explore different possible explanations for research misconduct (especially falsification and fabrication), and investigate whether they are compatible. We suggest that to explain research misconduct, we should pay attention to three factors: (1) the beliefs and desires of the misconductor, (2) contextual affordances, (3) and unconscious biases or influences. We draw on the three different narratives (individual, institutional, system of science) of research misconduct as proposed by Sovacool to review six different explanations. Four theories start from the individual: Rational Choice theory, Bad Apple theory, General Strain Theory and Prospect Theory. Organizational Justice Theory focuses on institutional factors, while New Public Management targets the system of science. For each theory, we illustrate the kinds of facts that must be known in order for explanations based on them to have minimal plausibility. We suggest that none can constitute a full explanation. Finally, we explore how the different possible explanations interrelate. We find that they are compatible, with the exception of explanations based on Rational Choice Theory and Prospect Theory respectively, which are incompatible with one another. For illustrative purposes we examine the case of Diederik Stapel.

Keywords Research misconduct · Explanations · Research integrity

1 Introduction

Over the past few years, interest in research misconduct has substantially increased (Gunsalus 2019). While not everyone agrees about what should be labeled a research *misbehavior*, there is general consensus on what has been called research *misconduct*: falsification, fabrication and plagiarism (FFP) (Lafollette 2000; Steneck 2006). This consensus is reflected in codes of conduct, both national and international (ECoC 2017; NCCRI 2018).

This paper has a twofold aim. First, to explore and discuss a number of possible *explanations* of research misconduct, and second to use this as a case study for the more

✉ René van Woudenberg
R.van.Woudenberg@vu.nl

¹ Department of Philosophy, Vrije Universiteit, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands

philosophical question: how do these different explanations relate to one another: are they compatible, or are they not?

This paper potentially has practical relevance in that explanations of research misconduct can be expected to give a handle on what can be done to prevent research misconduct. This being said, this paper focuses on explanation, not prevention.

The paper is organized as follows. In Sect. 1 we describe various types of research misconduct, and describe one actual case for concreteness' sake, as well as for the sake of future reference. Section 2 discusses what to expect from an explanation. The next section presents and discusses a number of explanations of research misconduct and explores what needs to be known if those explanations are to have some minimum level of credibility. In Sect. 4 we discuss the more philosophical question of how these explanations hang together. We conclude with some overall remarks.

2 Research Misconduct

The most extreme kinds of research misbehaviors—fabrication, falsification, and plagiarism (FFP)—are at the same time not the most frequent ones (Martinson et al. 2005; Fanelli 2009). Much more frequent are the numerous 'minor offences', the many cases of 'sloppy science', the 'questionable research practices' (QRPs) (Steneck 2006). According to recent surveys, examples of frequent QRPs are: failing to report all dependent measures that are relevant for a finding (Fiedler and Schwarz 2016), insufficient supervision of junior co-workers (Haven et al. 2019); selective citing to enhance one's own findings or conviction; and not publishing a 'negative' study (Bouter et al. 2016; Maggio et al. 2019). Despite their presumed frequency, assessment of the wrongness of the QRPs can be less than straightforward. Here, context, extent and frequency matter. The wrongness of FFP is more evident and codes of conduct are typically developed in order to prevent these. (For an excellent overview of different reasons for using a wide or narrow concept of research misconduct, see Faria 2018).

The reason why research misconduct needs to be prevented is somewhat different for falsification and fabrication compared to plagiarism. Whereas falsification and fabrication distort the *creation* of scientific knowledge, plagiarism need not distort the field nor hamper its progress. Plagiarism fails to connect the knowledge to its proper origin, but it need not distort scientific knowledge per se (Steneck 2006; Fanelli 2009). Also, explanations for plagiarism can be expected to differ from explanations for falsification and fabrication. Some plagiarism, for example, is committed by non-fluent English authors who borrow well-written sentences or even entire paragraphs for their own work, which is an explanation that is not available for cases of falsification and fabrication. We therefore focus on the latter two.

For illustrative purposes, we will examine a case of actual research misconduct in order to review the applicability of explanatory theories of research misconduct. We chose the case of Diederik Stapel for two main reasons. First, because his fraud has been established beyond reasonable doubt. Second, because there is sufficient publicly available information about the case: information about the committees' way of assessing the case, as well as about Stapel's own responses and reflections on his case. The more details of a case that are available, the better we can discuss the explanatory power of the theories we shall review. With the disclaimer that it is not our aim to provide an explanation of Stapel's

fraudulent behavior, and that others have produced interesting accounts of it (for example, see Abma 2013; Zwart 2017), we now offer a very brief description of the Stapel case.

Diederik Stapel was a professor of cognitive and social psychology. His research included topics such as the influence of power on morality, the influence of stereotyping and advertisements on self-perception and performance, and other eye-catching topics (Vogel 2011). He was an established figure whose findings often appeared in (inter-) national newspapers. Stapel was accused of data falsification by three whistle blowers from within Tilburg University, where Stapel was employed in 2011, the year the case became public. In total, three committees investigated whether Stapel's work at the University of Amsterdam, University of Groningen and finally University of Tilburg, was indeed fraudulent (Levelt Committee, Noort Committee, Drenth Committee 2012). The committees established that, whilst the studies were carefully designed in consultation with collaborators, Stapel fabricated the data sets from scratch. In another variant, the data were gathered but altered by Stapel after a student-assistant had forwarded them to him. Finally, Stapel had at times reached out to colleagues inviting them to use some data he claimed to have 'lying around'.

Stapel has admitted that he engaged in these practices. The committees concluded that Stapel intentionally falsified and fabricated data. None of Stapel's co-authors were found to have collaborated with him in this regard. We will provide more information about the case as we proceed.

3 What to Expect From an Explanation

It is fair to say that currently, we have no single unifying theory of explanation (Woodward starts his book with a similar remark, see Woodward (2003)). What we have is a wide assortment of ideas that are all claimed to be at least sometimes relevant for understanding explanation. One idea is that explanation is closely linked with *causation*: an explanation of X can be achieved by pinpointing the causal factors relevant to X. Another that it is closely linked to *laws*: an explanation of X is achieved by referring to laws under which X can be subsumed. Yet another idea is that explanation is linked with *unification*: an explanation of the phenomena X, Y and Z is achieved by showing that X, Y and Z are special cases of a more general phenomenon GP. A further idea is that explanation sometimes has to do with *reasons* (as opposed to causes): an explanation of a person's action A is achieved by citing her reasons, i.e. her beliefs and desires, for doing A.¹ In the social and behavioural sciences, this idea is sometimes coupled with the idea mentioned above that explanation is linked with laws. This approach to explaining human behaviour aims to formulate empirical generalizations of the form: If person P desires D, and believes that action A is the most efficient means of attaining D, then P does A. The hope is that such generalizations can be improved so as to state genuine laws, laws that enable prediction. Whether this hope is a

¹ We note that reasons can serve different roles: they can be *motivating* and they can, even at the same time, be *normative*. P's motivating reasons are the reasons *for which* P did A—the *considerations in light of which* P did what she did, and that motivated her for doing A. Normative reasons are the reasons that P would cite in favor of her action A, reasons that would show that A was the sensible, or right thing to do. This way of making this distinction is borrowed from Dancy (2000). Anscombe (2005) offers a subtle analysis of the notion "explaining behavior".

realistic one need not detain us here. The important point to note is that reference to a person's reasons often has explanatory force.

However, it is often not *just* a person's reasons that have explanatory force; they often have it in conjunction with what we shall call "affordances": the specific situations in which a person acted and in which certain possibilities are open to him. The explanation of the fact that A shot B cannot consist of merely citing A's desire that B be dead and his belief that pulling the trigger was a way to attain that goal. A factor in the explanation should surely be the availability of a gun to A. The availability of the gun is a contextual affordance for A.

We should add that some behaviors can be explained independently of the actor's reasons, and independently even of the actor's being aware of displaying those behaviors. There are unconscious influences on human behavior, like the biases and heuristics that psychologists have been researching, and reference to them can also do explanatory work (see Gilovich 1991; Kahneman 2011).

To conclude: if we want to explain cases of research misconduct, we should pay attention, among possible others, to the following factors:

- I: the desires and beliefs of the misconductor, meaning his or her (motivating) reasons;
- II: the contextual affordances available to the misconductor;
- III: unconscious influences²

In an actual case of misconduct, all these factors may be at work. We should therefore heed the distinction between *partial* and *full* explanations. A full explanation of an event specifies *all* the factors that jointly guarantee the occurrence of the event. A partial explanation, by contrast, specifies a factor, or several factors, that facilitate the occurrence of the event, but do not guarantee it. It remains an open question (for us at least) whether full explanations of human behavior are even possible.

Explanations in the social sciences can take various forms. One that will figure quite prominently in our discussion are inferences to the best explanation (IBEs). A key feature of IBEs is that the factor doing explanatory work is not directly observed, but concluded to.³

4 Explanations of Research Misconduct

In a helpful article, Benjamin Sovacool (2008) distinguishes three 'narratives' about research misconduct: one in terms of (1) impure individuals, another in terms of the (2) failures of this-or-that particular university or research institute, and yet another in terms of (3) the corrupting structure of the practice of modern science as such—three narratives that he suggests are incommensurable. Even if these narratives do not explain in

² Note that the factors we describe seem to match up with what has been termed levels of explanation, e.g. an explanation using desires and beliefs would be an explanation on the personal level, etc. (see Owens 1989). Yet, we will not focus on the question of whether an explanation on one level is more fundamental than an explanation on another; our aim is merely to assess the plausibility of the explanations and whether they are compatible with each other.

³ See Lipton (2008, Ch. 4). Standard examples of IBEs are the doctor's inference that his patient has measles, since this is the best explanation of the symptoms; and the astronomer's inference to the existence and motion of Neptune, since that is the best explanation of the observed perturbations of Uranus.

any straightforward way individual cases of research misconduct, they are helpful for two reasons.

First, narratives can deliver cognitive goods that are distinct from explanations—they can provide understanding. And, as Peter Lipton (2009) has argued, there can be understanding without explanation. Even if we have no explanation of Stapel's fraudulent behavior, it does give insight into the whole affair if the evidence indicates that Stapel was only one bad apple, or if it indicates that the institute at which he worked was failing in important respects, or if the whole structure of science turns out to be corruptive. Second, Sovacool's narratives are helpful as they *do* point to places we could look for explanations. For example, the narrative that a case of research misconduct is due to an impure individual (and not a failing research institute, nor something like the corruptive structure of science as such) does not explain in any detail why Stapel engaged in the misbehavior he did, but the narrative (if true) *does* point to what is needed for such an explanation: the nature of his impure character needs to be understood, so that we can see how Stapel's specific impurity led to the misbehaviors that made him notorious. Likewise, the narrative that the misconduct is due to a failing research institute does not explain Stapel's behavior, but it *does* point (if true) to where to look for an explanation: to the operative rules and procedures of the institute, perhaps, to its 'culture' or 'climate' ('there was an atmosphere of terror'), etc.

Of course, things get complicated here. For if Stapel's misbehaviors are due exclusively to factors covered in the narratives about the institutions he was part of (or about the structure of science as such), then we should expect other members of those institutions to have displayed similar misbehaviors—which, as far as we know, they have not. And this is a reason for thinking that Stapel's misbehaviors are due *not* exclusively to institutional failings, but also, say, to personal impurities like character flaws. The distinction between partial and full explanations is a recognition of this complication.

We draw attention to the fact that whereas explanations under Sovacool's first narrative will typically refer to type I and III factors (beliefs and desires; unconscious influences), explanations under Sovacool's second and third narrative will refer to type II factors (contextual affordances). Since all these factors, possibly and likely, can play a role in cases of research misconduct, we need not assume that the explanations under Sovacool's three narratives are per se incommensurable if that entails they are incompatible. In fact, as we will argue in Sect. 4, most of these explanations are compatible with each other, as they are partial at best.

To conclude: Sovacool's narratives do not offer explanations of cases of research misconduct, but they point to where to look for explanations. We discuss six⁴ different (types of) theories that might help explain research misconduct.⁵ Our aim here is to specify what

⁴ Our search for theories was guided by a similar endeavor of Gjalte de Graaf's (2007), in which he discusses a number of theories that purport to explain corrupt or fraudulent behavior in public administration, such as taking bribes. We supplemented his list with additional theories where relevant. Interestingly, de Graaf also included correlation 'theories', but as these are not theories in that they do not contain an idea about the explanatory mechanism, we do not review them in-depth but will elaborate briefly on their relevance in the concluding section.

⁵ Note that the (types of) theories De Graaf (2007) reviewed are theories of human behavior that come from different fields such as economics, sociology or criminology and seem to work on different levels. Hence, these theories often apply to terrains beyond fraud in public administration and may not even have been designed to explain fraud in public administration, but *have* been *appealed to* to explain fraud in public administration. In a similar vein, we explore whether these theories that *have* been coined to explain research misconduct are actually applicable and compatible.

we need to know about a specific case in order for such explanations to get a good start. Whether they are credible, is a further issue. We begin with four theories that fall under Sovacool's first kind of narrative.

4.1 First: Rational Choice Theory

Sometimes labeled 'rational choice theory', this theory has its origins in economics. It starts from an individual that is portrayed as rationally considering different options to tackle a particular problem. Rational Choice says that an individual actor faced with a risky outcome selects the specific behavioral action that yields the maximized anticipated pay-offs, where the utility of his behavior is weighted by the probabilities of its occurrence. The domain of the utility function is absolute benefits and costs. The individual weighs the costs and benefits attached to each option, and next makes the calculation, on the basis of which she makes a decision.⁶ This theory, that refers to type I factors only (beliefs and desires), is appealed to in the research integrity literature by Wible (1992) as well as by Lacetera and Zirulia (2011).

Suppose we apply this theory to Stapel's case. We will first describe what we think needs to be the case if this theory is going to provide an adequate explanation of his misconduct. Next we discuss whether (we know) these things are indeed the case.

If this theory is to explain Stapel's misconduct, we should envisage Stapel as a rational agent⁷ who is calculating the costs and benefits, i.e. the utility, of cheating compared to playing it fair (i.e. observing the rules and principles that we now find in the numerous Codes of Responsible Conduct of Research). The benefits of (undetected) cheating probably include: more publications (or: more publications with outcomes that would be considered remarkable), which would contribute to greater prestige, which would increase the chances of obtaining more research funds, which would mean gaining more visibility, power and influence. The costs of cheating probably include: the fear of being found out (and fear of whatever else is set in motion by it: retraction of publications, loss of research funds, loss of prestige, loss of job, etc.), which means that one must always be on one's guard; loss of self-respect; not contributing to the (great!) cause of science. The costs of playing it fair include, probably: often having research results that are not significant and/

⁶ We are aware that rational choice theory is sometimes used as a general paradigm that is not to be applied to individual cases because the notion of a "rational choice" is deemed to be no more than a useful theoretical fiction. We side with those authors that have used rational choice theory to shed light on individual cases of human behavior.

⁷ What does it mean to be a 'rational' agent in this case? In jurisprudence, an important consideration for holding someone accountable is whether that person had the right mentality, or *mens rea*. The four generally distinguished levels of mentality are purpose, knowledge, recklessness and negligence. Each of these correspond to a different extent to which the researcher, in our case Stapel, could be held accountable for his deeds, with the first being the highest level of accountability, purpose. Stapel's case maps most closely onto this level—in his own writings, he is explicit about his intention to deceive others. The mental state of knowledge would look something like this: a colleague of Stapel had reasonably strong doubts about Stapel's conduct, but decided to work with him regardless. Recklessness could perhaps be applied to cases of falsification, where a researcher runs a data analysis she does not fully understand and finds a significant result that she reports regardless. The level of negligence does not seem to work in our case, as it is unlikely to engage in misconduct out of negligence. So if one applies rational choice theory to cases of misconduct, one should be clear about whether that conduct was intentional, with knowledge, reckless or negligent, as rational choice theory seems more apt to explain cases where the trespasser had a *mens rea* of purpose or knowledge, compared to negligence or recklessness.

or interesting, which decreases the likelihood of one's research being published, which decreases the chances of getting research funds, and of making an impact. The benefits of playing it fair include: doing what one, from a moral point of view, ought to do, behaving in a responsible way (and virtue is its own reward, as the proverbial wisdom has it); increasing the chance that you will have research results that are genuine contributions to the cause of science; increasing the chance of receiving recognition that is based on substance.

If rational choice theory is going to give an adequate explanation of the falsifications and fabrications committed by Stapel, he must have engaged in a cost/benefit analysis of the cheating option as compared to the playing fairly option—and on that basis have decided that falsification and fabrication 'pay'.

Is there any evidence that Stapel did engage in a cost/benefit analysis of this sort? There are two main types of possible evidence here: the misconduct investigation reports and Stapel's own accounts. From the report (Levelt et al. 2012) on Stapel's misconduct, we could deduce that the costs—at least, the fear of being found out—seemed low: "It was easy for researchers to go their own way. Nobody looked over their shoulder..." (p. 39). Stapel's own account⁸ also points in this direction: "So when I started to fake my studies, it was very, very easy. I was always alone. Nobody checked my work; everyone trusted me. I did everything myself. Everything. I thought up the idea, I did the experiment, and I evaluated the results. I was judge, jury, and executioner. I did everything myself because I wanted to be one of the elite, to create great things and score lots of major publications" (p. 118–119).

Yet, it remains somewhat questionable whether Stapel actually engaged in a cost/benefit analysis. But this does not mean that the rational choice theory explanation is false or wrong. Stapel's engaging in such an analysis is at least a *possible* outcome of a rational choice IBE, for his fabrications and falsifications *may* be best explained by his having made a cost/benefit analysis. Whether it indeed *is* the *best* explanation, depends, of course, on the strength of alternative explanations. Moreover, as we noted, explanations can be *partial*. Rational choice theory, then, may offer only a *part* of a full (or fuller) explanation. As a matter of fact, this IBE, even if it is correct, can at best be a partial explanation only. For, as we suggested in the previous section, there must be contextual affordances (so type II factors), in this case: structures and systems that allow for the possibility of falsification and fabrication. And these affordances fall outside the scope of rational choice theory, as do type III factors.

4.2 Second: Bad Apple Theories

Like rational choice theory, this theory too has its roots in economics. Here, the individual is depicted as someone with a flawed (moral) character. This flawed character is subsequently causally linked to corrupt acts. Greed is sometimes deemed to be an element in a flawed character. An example of a full-scale faulty character has been labelled in the

⁸ Note that this regards the translation of Stapel's 2012 autobiographical book by Brown (2014). Caution is needed when interpreting these statements, as it is arguably *oratio pro domo*. Stapel seems to acknowledge this as his foreword reads: "This is my own, personal, selective, biased story about my downfall." (p. iii). Similar to Zwart (2017), it is not our primary concern whether "the autobiographical account actually corresponds with the facts... but rather what can be learned and gained from this ego-document" (pp. 211–212).

literature as the Machiavellian personality type, that deems that the prestige associated with a particular goal justifies any means to attain it, even if those would be seen as unethical. Hren et al. (2006) studied Machiavellism in relation to moral reasoning and Tjldink et al. linked personality types such as a Machiavellian character to research misbehaviour (Tjldink et al. 2016). Bad apple theories refer to type I factors only—to reasons that motivate certain characters to behave in certain ways.

If we apply this theory to Stapel and ask what should be the case if bad apple theories are to provide an adequate explanation of his misconduct, it is clear that he needs to have, or at the time have had, a flawed moral character—he needs to have a Machiavellian personality type for example, or some other flawed moral character.⁹

Is there evidence that Stapel had a flawed moral character at the time—evidence coming from psychologists and psychiatrists, for example, who have done something like a personality-analysis on him? The only evidence that would point in that direction appears in Stapel's own book (Stapel 2014): "*It takes strong legs to carry the weight of all that success. My legs were too weak. I slipped to the floor, while others—maybe wobbling, maybe with a stick to lean on—managed to stay upright. I wanted to do everything, to be good at everything. I wasn't content with my averageness; I shut myself away, suppressed my emotions, pushed my morality to one side, and got drunk on success and the desire for answers and solutions.*" (p. 148, emphasis original). Yet, this one passage seems insufficient as a basis for a solid psychological verdict on his character, and as far as we know we have nothing else to go on that is publicly available and would reliably demonstrate a flawed character.

Note that when we refer to a flawed character, we do not mean to insinuate that Stapel had no moral awareness whatsoever. The report (Levelt et al. 2012) on his misconduct explicitly mentions that he taught the research ethics course. Stapel's account (Stapel 2014) confirms this: "I'm the teacher for the research ethics course, in which I get to discuss all the dilemmas with which I'm confronted every day, and for which I always make the wrong choice." (p. 129).

Even if we have no solid basis to draw a conclusion about Stapel's moral character, this doesn't mean a bad apple explanation can be ruled out. For it is possible to make an IBE, based on a bad apple theory, to the effect that Stapel's fraudulent conduct is best explained by the fact that he had, at the time, a flawed (moral) character. Whether this is really the best explanation, depends, again, on the strength of the available alternatives.

It seems clear, however, that bad apple theories, even insofar as they are correct, cannot give us a full explanation of Stapel's misconduct. For there must be contextual affordances that allow flawed moral characters to commit acts of fabrication and falsification—and these are part of a full(er) explanation of the misconducts at hand.

4.3 Three: General Strain Theory

Another theory that could be headed under the individual narrative is General Strain Theory (henceforth: GST) as originally developed by Agnew (1992) who worked in the sociology of crime. GST sees misconduct as originating in stress or strain. These states of stress and strain bring about a negative emotional state in the researcher, like anger, sadness or

⁹ Although we mostly discuss moral character flaws, it seems plausible that intellectual character flaws, such as insouciance, play a similar role—insouciance example taken from Cassam (1992).

depression—which are, broadly speaking, type I factors. As a third step, GST posits that the behavioral strategies researchers adhere to in order to cope with these negative states differ, and, importantly, strategies may include deviant behavior (in our case: research misconduct). This theory, which has been coined as playing a role in explaining research misbehavior by Martinson et al. (2010), is put forth in the Institute of Medicine’s report *Fostering Integrity in Research* (NASEM 2017), and recently came forward in research by Holtfreter et al. (2019) wherein they asked US scientists what factors they believed to play a role in research misconduct.

If this theory is to do explanatory work, we need to know whether Stapel faced prolonged stressful situations, so prolonged that they put him in a persistent negative state. The report on Stapel’s misconduct is silent on this issue. In his book, Stapel himself, though, talks of a persistent state of stress he experienced: “Nothing relaxes me any more... but I feel stressed and restless. I want everything, and everything has to happen now. I want out. I don’t want to have to write papers any more. I want to start over, get away from this fantasy world I’ve created, get out of this system of lies and half-truths, to another city, another job” (Stapel 2014, 131). However, he experienced this *after* he got into the habit of altering his data.

GST theory presupposes that behavioral strategies to cope with the negative emotional states differ. Thus, whereas Stapel’s colleagues facing similar strains found other ways to cope, he turned to deviant behavior. But this is also a caveat: What exactly made Stapel turn to deviant strategies? Perhaps his environment was crucially different in some way, which fueled his urge to create spectacular results? In any case, GST can thus, at best, be a partial explanation. That is not to say that GST can be ruled out entirely, as it is possible, via an IBE, that his misconduct could be explained by GST—whether that is also the best explanation depends on the explanatory force of the alternative theories.

4.4 Four: Prospect Theory

The final theory that we shall consider under Sovacool’s first narrative is prospect theory. The roots of prospect theory lie in the psychology of risk, but the theory has also been used in behavioral economics. In their study of risky choice, Kahneman and Tversky (1979; Kahneman 2003) found that individuals are more strongly motivated by fear of loss than potential gain, and are inclined to avoid risk when faced with potential gains, yet seek risk when faced with potential losses. Bearing in mind that the reference point of the individual researcher matters (their context—whether that is one in which the researcher is faced with potential losses or gains), prospect theory would predict that researchers faced with potentially losing their job, tenure or other meaningful resources would be more prone to take risks, or in our case, to engage in research misconduct, than colleagues who face no such threats. This theory refers to type I and II factors, as the behavioral tendencies involved may, but need not, go unnoticed by the subject. The National Academies’ report *Fostering Integrity in Research* offers this as a possible explanation in its chapter on the causes of deviance (NASEM 2017).

For this theory to explain Stapel’s deeds, we need to know whether, at the point in time when he falsified or fabricated datasets, he was faced with the threat of losing his job, or tenure, or other meaningful resources. In addition, it would be useful to know if the opposite situation occurred, where Stapel was faced with a potential gain, perhaps greater chance of having his research accepted in a high-impact journal through the risky behavior of falsifying his data, and decided against it.

Stapel's book contains a passage of his reflection that reads: "There was no pressure, no power politics, no need to produce patents or pills, to compete in the marketplace or make a pile of money. It was always purely academic, scientific research, which makes any form of cheating even harder to understand." (2014, 188). Another passage seems to point more at the potential for gain as a driving force: "I couldn't resist the temptation to go a step further. I wanted it so badly. I wanted to belong, to be part of the action, to score. I really, really wanted to be really, really good. I wanted to be published in the best journals and speak in the largest room at conferences." (p. 102–103).

The report (Levelt et al. 2012) does not provide direct information on these issues, but it does detail that 55 of Stapel's publications rested on falsified or fabricated data. Even if we put aside the idea that different papers can be based on the same dataset, how often can one be faced with potentially losing their job, tenure or another meaningful resource? It seems likely that there were other factors at play, too. Again, that is not to say that prospect theory cannot be an explanation for research misconduct, but that it can at best be a partial explanation. And even if, in Stapel's case, there was no direct evidence that he feared losing his job, this potential threat could be inferred via an IBE. This in turn sparks the question whether it is also the *best* explanation, given its competitors.

We now move on to consider a theory that aims to explain misconduct by referring to the institutions and organizations in which the perpetrator works, and thereafter to a theory that aims to explain it by referring to the structure of the practice of modern science in general. Explanations based on these theories refer to type II factors, contextual affordances.

4.5 Five: Organizational Culture Theories

These theories find their roots in organizational psychology. They have in common that they consider people as working in an organization with a specific culture and a particular structure, and argue that these have an effect on individuals and their behavior. An assumption underlying these theories is that there is a causal path from a certain organizational culture, to a particular mental state, to an individual's behavior.

One particular organizational culture theory, called organizational justice theory, is based on the idea that people who perceive themselves to be treated fairly by their organization,¹⁰ behave more fairly themselves. Conversely, when the organizational procedures are perceived as unfair, people are more likely to engage in acts that make up for the perceived unfairness, e.g. falsifying or fabricating their data. Martinson and colleagues (Vries et al. 2006; Crain et al. 2013) have investigated this theory and they report that researchers who perceived their treatment as unfair were more likely to engage in research misconduct.

There are various ways in which the organization can influence the behavior of researchers, and the organization itself is not immune to external influences.¹¹ The

¹⁰ The organization is often studied through the organizational culture (the values, beliefs, and norms that help shape members' behavior) and the organizational climate, defined as "the shared meaning organizational members attach to the events, policies, practices, and procedures they experience and the behaviors they see being rewarded, supported, and expected" (Schneider et al. 2013, 115). We will look at both in our consideration of organizational justice theory, but as policies and procedures are more observable than values and beliefs, we will focus more heavily on the former when reviewing the empirical materials available.

¹¹ It can be hard, in the case of academic research, to pinpoint the boundary at which the culture ends and the outside begins, which can be seen as a caveat of applying organizational justice theory to research misconduct. Sometimes we speak of the research culture in, say, psychology, referring to the scientific field at

Institute of Medicine's (IOM) report *Integrity in Scientific Research: Creating an Environment that Promotes Responsible Conduct* (IOM 2002) conceptualized the research organization as an open systems model. Within the organizational structure itself, there are policies and procedures in place that influence researchers, and within the organizational processes the IOM report emphasizes the role of leadership and supervision. These last two are especially important, as studies on the organizational climate in academic and other settings found that organizational leadership, ethics debates and ethical supervision were associated with an ethical climate. The system is open in that it produces various outputs in the form of papers and other research related activities that in turn influence organizational inputs through funding and human resources, which in turn influence the organization again.

Another idea is that the organisational dynamics themselves can take such a form that everyone in the organization begins to engage in questionable practices. This type of unethical conduct may then become so frequent that it slowly becomes the normal way of conducting research.

If we apply this theory to the misconduct of Stapel and ask what should be the case if his misconduct is to be adequately explained by it, we must say that the culture and structure of the organizations he was employed by, somehow induced his conduct. Either there should be indications that he was mistreated by his organizations or there should be evidence that his work environment was perverted altogether. Delving deeper: Is there information available on their policies, the degree to which leadership emphasized integrity, or whether open debates about integrity issues were a regular occurrence? There must, perhaps, have been reward systems in place that triggered misconduct, or some element of an organization's culture that did the trick.

So, if such an explanation is to work for the Stapel case, what we need is insight into the culture and structure of the organizations that he worked with. Stapel seems to believe that culture played a role (Stapel 2014, 171): "I'm not the only bad guy, there's a lot more going on, and I've been just a small part of a culture that makes bad things possible." Even if there was no direct evidence available about Stapel's research culture, it might be possible to make an IBE here too: from his misconduct we can draw conclusions suggesting a bad organizational culture and bad organisational structures—the latter explaining the occurrence of the former.

Interestingly, the report (Levelt et al. 2012) about Stapel's misconduct devotes an entire chapter to the culture in which his fraud took place. It is described as "a culture in which scientific integrity is not held in high esteem" (p. 33) and "even in the absence of fraud in the strict sense, there was a general culture of careless, selective and uncritical handling of research and data." (p. 47). This may prompt one to believe that the culture indeed played a role in fostering Stapel's fraudulent behavior. However, the report (Levelt et al. 2012) presents culture as an explanation for why the fraud could *sustain* for so long—"The Committees are of the opinion that this culture partly explains why the fraud was not detected earlier." (p. 47)—not as one that brought about the fraud. Of course, this does not preclude

Footnote 11 (continued)

large. Related, internal means of promotion or tenure are influenced by review committees of papers and grants, which would traditionally be placed outside of the organizational culture (see also Martinson et al. 2010). Nevertheless, it seems reasonable to suppose that an individual researcher is most profoundly influenced by their local climate—by the policies that directly apply to them and by the practices they see their colleagues engage in and be rewarded for.

the organizational culture from being a potential explanatory factor in the origination of the misconduct as well.

Are there indications that Stapel was structurally undervalued by his respective organizations, and treated unfairly? The report's (Levelt et al. 2012) information points in the opposite direction: "These more detailed local descriptions also reveal Mr Stapel's considerably powerful position, at any rate within the University of Groningen and even more so within Tilburg University. At the University of Amsterdam he already enjoyed a reputation as a 'golden boy'." (p. 38). To our knowledge, there is no public evidence of a culture that treated researchers unfairly or that suggests Stapel's deeds could be interpreted as a means to make up for perceived unfairness done unto him.

Can we know enough about the organization's culture and the structures of the units Stapel belonged to? Perhaps we can. But even if we do, the organizational culture explanation can at best be a *partial* one. For many other individuals who worked in the same organization, have not (we assume this to be so) committed acts of fabrication and falsification. For this reason we may think of an organization's climate and structure as contextual affordances that do not forestall misconduct, and do not cause it either, but do enable it.

Until a certain stage of investigation, it is possible to propose an organizational culture explanation of Stapel's behavior, namely as long as we have no evidence that any of the other explanations even partly explain it. At a later stage of the investigation, however, it should be possible to have more direct access to the organizational culture, as it should in principle be observable.

4.6 Six¹²: Ethos of Public Administration

Ethos of public administration theories, at times labelled Taylorism or New Public Management (NPM) theories, have their roots in economics, and, applied to research misconduct, fall under Sovacool's third kind of narrative. These theories center around a complex set of ideas and concepts: specialization, command, unity, efficiency and atomism. The ideas that connect these concepts are, firstly, that individuals are naturally isolated from one another and that only an organization, through a chain of command and a sense of mission, can unify individuals into a single, efficient and rational working unit. The second is that individuals tend to laziness, selfishness and are not interested in any social good beyond their own individual good, and that therefore organizational unity and discipline must always be maintained.

The perverting influences of NPM or Taylorism on the academic system can be expressed through different phenomena that Halfman and De Radder (2015) eloquently captured in their *Academic Manifesto*. They describe, among other phenomena, the "measurability for accountability" (p. 167), meaning the obsession with output quantifiers, be it publication indices, metrics, or impact factors. They also elaborate on the "permanent

¹² De Graaf's (2007) fourth type of theory is the theory of clashing moral values. The idea builds on Sellin (1983) and is that particular values that are held in high regard in the private atmosphere may lead to behaviors that are undesirable in the public or work atmosphere. Davis (2003) applied this to research misconduct cases: Take a researcher that comes from a culture where scientific productivity is the holy grail. After working for some time in a culture where adherence to ethical practice is regarded pivotal, some of the researcher's behaviors may be regarded fraudulent. Davis' argues that this can be fixed by subjecting the researcher to ethical training or developing codes of conduct that are endorsed widely. If this theory is to stand ground, Stapel should have been (successfully) socialized in a culture that held values which clash with ethical practice in extremely high regard. We think the analogy to research misconduct does not work here. Which value in the private atmosphere is supposed to do the explanatory work? If we look at the example from Davis, it is far from obvious that adherence to ethical practice clashes with productivity. We regard clashing moral values a nonstarter in the case of research misconduct.

competition under the pretense of ‘quality’” (p. 168), referring to the ‘hypercompetition’ where researchers compete against each other for funding in a ‘winner takes it all’ system, where it is the junior staff that do the bulk of the work, faced with temporary contracts and poor career opportunities (Halfman and Radder 2015).

Now, this extreme emphasis on effectiveness and performance can come at the cost of neglecting ethical issues and crowding out the values that motivate professional behavior and institute the organization’s mission. When this happens, it can lead to corrupt individuals. Overman et al. (2016) seem to subscribe to this proposition when they write: “Academic misconduct is considered to be the logical behavioral consequence of output-oriented management practices, based on performance incentives.” (p. 1140).

If this theory is going to explain Stapel’s misconduct, what should be the case is that he worked in an organization with a strong focus on performance and output in a way that crowds out values and the acknowledgement thereof. Perhaps he started out with an intrinsic desire to do good research. However, the more his work’s merit was determined by performance indicators and the more the focus was put on effectiveness, the more this intrinsic motivation was replaced by a desire to do good *according to* these performance indicators—to be effective and publish lots of papers. In addition, the emphasis on these performance incentives shifted attention away from responsible conduct of research.

Is there evidence that Stapel worked in such a system? Overman and colleagues describe that performance indicators indeed have become more evident among academic institutions in The Netherlands (they draw on research by Teelken (2015)). Do we have evidence that increased emphasis on performance accounts for Stapel’s actions? His own account acknowledges the pressures in contemporary science: “Science is an ongoing conflict of interests. Scientists are ... all in competition with each other to try and produce as much knowledge as possible in as short a time, and with as little money, as possible, and they try to achieve this goal by all means possible. They form partnerships with business, enter the commercial research market, and collect patents, publications, theses, subsidies, and prizes.” (Stapel 2014, 189–190).

Perhaps we should consider the role of these performance indicators plus the reality of hypercompetition as biasing Stapel’s view on research. Under their influence, he unconsciously focused more and more on effectiveness at the expense of ethical conduct. At some point, effectiveness itself became his main desire. One is reminded of Goodheart’s law: “When a measure becomes a target, it ceases to be a good measure”.

However, we are again left with the question why these indicators biased Stapel towards extreme efficiency and not his peers. Maybe his affordances were different from those of his peers, but these fall outside the scope of this theory. Hence the ethos of public administration or NPM, even if it is an acceptable explanation of misconduct, can best be thought of as a *partial* explanation.

As with the other theories, even if (so far) there is no direct evidence that a case of scientific fraud was caused (at least in part) by excessive emphasis on effectiveness and performance indicators, excessive emphasis could, indirectly, be inferred via an IBE. In which case the question arises whether it is also the *best* explanation, given its competitors.

5 Are the Different Explanations of Research Misconduct Compatible?

Having discussed six explanations of research misconduct, and having explicated what, for each of them, needs to be the case if they are to be accurate, if only partial, explanations, we now address the second question that we have for this paper: how do these explanations

relate to each other? Two different explanations of the same phenomenon, E1 and E2, can be compatible, or they can be incompatible. And if they are compatible, further qualifications can be added—for example that E1 and E2 “add up”, or that they reinforce each other, or that one weakens the other. We will see examples below of each of these sorts of relationships.

Given that we have six explanations on our hands, this means there are 15 pairs of explanations to consider. We can reduce this number because each of the four explanations under the first narrative are in their nature compatible with the explanations under the second (institutional) and third (system of science) narratives. This is in the nature of the case, as the first focus on qualities of the miscondactor, and the latter two on contextual affordances—none of which, we suggested, constitute full explanations. We do not want to make this point only at this abstract level, but want to offer one illustration. Consider bad apple explanations and organizational culture explanations. It would seem that such explanations (of the same behavior) are *at least* compatible. If cheating can be adequately, if only partly, explained by reference to the ill treatment that the cheater has suffered in an earlier stage, then this explanation can be augmented by the additional explanation that the cheater has a failed moral character. And if cheating can adequately, if only partly, be explained by reference to the culture within the organization that the cheater worked with, then this explanation can be augmented by the additional explanation that the cheater has a failed moral character. So these explanations are *at least* compatible. *At least*, for it is possible (and plausible) that these explanations reinforce each other in this way: failed moral characters will tend to make organizational cultures bad, and bad organizational culture will tend to make moral characters fail. Failed moral characters in organizations with a bad culture, will tend to feel at home like fish in water. Applied to Stapel: the explanation of his misconduct can be explained by reference to his failed moral character (to akrasia perhaps), but also by reference to the culture of the organizations with which he worked. And the two explanations can reinforce each other, as bad characters breed bad cultures, and bad cultures breed bad characters.

As is in the nature of the case, the explanations under the second and third narratives, being Organizational Culture explanations and NPM explanations, are compatible as well. This point can also be made in a more concrete way. Since NPM will foster a particular kind of culture within an organization, and since a particular kind of culture will be especially sensitive to the down-sides of NPM, explanations of misconduct that refer to culture and to NPM are compatible, and they even reinforce each other. Applied to Stapel's case, his misconduct can be explained, partly, by reference to organizational culture, and this can be augmented (and so make for a more complete explanation) by reference to the down-sides of NPM—and these two reinforce each other.

Since the first narrative covers four explanations, there are six pairs to check for compatibility. The *first* pair we consider is Rational Choice explanations and Bad Apple explanations. We may feel pulled in two directions here. Suppose someone is a bad apple, i.e. displays a defective moral character (perhaps the person suffers from akrasia), then we may think that his choice can never be rational, because his defective moral character prevents him from making such a choice. On the other hand, if making a rational choice consists of weighing the costs and benefits of an action as compared to alternative actions, then it would seem that someone with a defective moral character can engage in rational choice making as well—even if the outcome of the calculation is not what we would like it to be. Since it is *formal* (“means-end”) rationality that rational choice theory works with, it seems that a rational choice explanation is compatible with a bad apple explanation of the same behaviour. Applied to the Stapel case: an explanation of his misconduct in terms of

character flaws (like *akrasia*) is compatible with the claim that his choice to cheat was the outcome of a rational cost–benefit analysis.

The second pair of explanations we consider is Rational Choice and General Strain. This pair puts before us the question whether strain and stress prevent a person from making a rational choice. On the face of it, stress and strain may lead a person to select a goal that he would not have selected in the absence of it; and given the goal, he may have calculated the means to attain it. Alternatively, a person may have set himself a goal, while stress and strain influence the calculation of the means to attain it. The influence may be that certain means become live options that were dead, or that options that were alive, die. But given the options, a stressed person may still make what he thinks is a fair calculation—fair not in a moral but in a formal sense. Either way, the explanations based on Rational Choice and General Strain are compatible. Applied to the Stapel case: stress and strain may have led him to set the goal of achieving high-profile publications, and rational choice deliberation suggested to him that fabrication and falsification were the ways to attain that goal. Alternatively, Stapel had set himself the goal of achieving high-profile publications, and strain and stress led him to calculate that fabrication and falsification were the best ways to attain the goal.

Third, Rational Choice and Prospect Theory, by contrast, do not deliver compatible explanations. For the former assumes that an actor will always seek maximal gains based on the probability of occurrence, while the latter says that fear of loss tends to be a much stronger motivator of behavior than the potential for gain, and also that individuals tend toward risk aversion when confronted with potential gains but bias toward risk seeking when confronted with avoiding potential losses.

Applied to Stapel: Rational Choice explains his fraudulent behavior by reference to a rational calculation he has made so as to have maximal gains, while Prospect Theory predicts that, given Stapel's stable job's situation (he had a tenured position with no fear of losing it), he would be less likely to make the risky choices that he did make.

The incompatibility should come as no surprise, as Prospect Theory was expressly developed as an alternative to Rational Choice (Thomas and Loughran 2014).

General Strain and Bad Apple approaches are compatible. If stress and strain induce deviant behavior, then it does so in virtuous persons and bad apples alike. Strain explanations and bad apple explanations are compatible, and they may even reinforce each other, in that it is plausible to think that bad apples make even worse choices if they also experience stress and strain—and that strained persons make worse choices if they have flawed moral characters. Applied to Stapel: if he had a flawed moral character, he may already have been open to cheating, but if he was also under stress and strain, then the cheating option may have become even more salient.

Prospect Theory and Bad Apple theory are also compatible. As indicated, Prospect Theory predicts that people faced with the prospect of losing their job or other meaningful resources, will be more inclined to take risks—and if this holds, it holds for bad and good apples alike. The two explanations of behavior it suggests, can both be correct, if only partially. Applied to our case: if we counterfactually assume that Stapel's position was at stake, and also that he had a flawed moral character, then both these factors can be referred to for explanatory purposes—and both explanations can be correct.

The sixth and final pair to consider is that of Prospect and General Strain. Strain and stress may be real in a person who, when faced with serious loss of meaningful resources, is more prone to take risks than when not so faced. Hence, two explanations of a person's behavior based on their own respective theories, can both be true and hence be compatible. However, if a person is experiencing stress and strain, while at the same time there

is no threat of loss of meaningful resources, then the two theories yield incompatible explanations. After all, Prospect Theory tells us that people are risk-averse. The impulse to deviant behavior generated by strain and stress would be mitigated by the impulse to risk-aversion. In that case we might say that the two theories are compatible—but that the explanations do not reinforce each other, nor do they add up, but rather the one weakens the other in the sense that the effect that one theory predicts does not occur to the degree it would have in the absence of the other effect. If we again assume that Stapel was experiencing stress and strain (which already motivated him towards deviant behavior) and he was also facing the threat of losing meaningful resources (which inclined him to take more risks than he would otherwise have taken), then the explanations reinforce each other. But if he was experiencing stress and strain, yet there was no fear of losing meaningful resources, then the theories lead us to expect deviant behavior to a lesser degree than if there was also a threat of loss.

6 Concluding Remarks

We have discussed six explanations of research misconduct, and how they relate to each other. We argued that most theories are compatible with each other, with the exception of Rational Choice and Prospect. Suppose now we concentrate on explanations that are compatible. Can we conclude that those pairs offer *full* explanations? For a number of reasons we cannot. First, we have only looked at pairs among the six theories we have discussed. But triplets of them may offer fuller explanations, and quartets of them even fuller. Second, there are explanations of research misconduct that we have not discussed, but that can be added to the fold.¹³ Third, a large body of research that investigates research misconduct takes the form of correlation ‘theories’ that map significant correlations between (some measure of) research misconduct and some other factor of interest. Of course, correlation does not equal causation. Take this one step further: on a narrow reading of theory—“an idea that is used to account for a situation”—it seems incorrect to speak of correlation *theories*. Correlations map temporal co-occurrence beyond some degree of doubt. The idea or link that is to explain this co-occurrence is often thought up post-hoc as a rationalization, but it is not (yet) a fully-fledged theory.

However, that does not render correlational research meaningless for explaining research misconduct. Similar to narratives, correlational research results deliver cognitive goods—they give knowledge about factors that in some way play into the misconduct. Along that same line of reasoning, they serve as a pointer for further theorizing that may at some point be formalized into a theory.

Still, we are left with the question whether it is sensible to suppose that, drawing on all correlational research and supplemented with the types of theories we reviewed here, one can *fully* explain research misconduct. There seem two avenues to take, both reconcilable with what we argued above, and these avenues are connected to one’s stance on

¹³ See for example Rajah-Kanagasabai and Roberts (2015) that use the theory of planned behavior to explain research misconduct in students. Because our review focused on misconduct among academic researchers, and it had not been coined outside the realm of students, we chose to not review the theory of planned behavior in-depth here. Alternatively, Hackett (1994) reviewed anomie as a possible explanation for researchers engaging in research misconduct, but he disregarded anomie so persuasively that we chose not review it here.

free will. Either one believes that humans are free and this will render some part of their behaviour—especially complex behaviors, like research misconduct—inexplicable. Or one believes that humans are not free and that scholars have not yet found the (final) key to the explanatory puzzle. It seems natural to think that this key, if it exists, is to be found somewhere along the lines of unconscious factors that influence human behavior, such as biases or heuristics. We tend to the first view.

A further point we would like to make is that although this paper is focused on theories coined to explain falsification and fabrication, these theories also seem relevant when explaining lesser trespasses, such as QRPs. In fact, for those QRPs that teeter on the edge of falsification—take *p*-hacking or HARKing (hypothesizing after results are known)—it seems natural to suspect that when we apply the theories reviewed here to explain the occurrence of those QRPs, we likely run into similar problems that we encountered when trying to explain research misconduct. And since explanations of research misbehavior—here encompassing both FFP and QRPs—feed into our ideas about prevention of research misbehavior, extending our theories and models on how to explain may help us to prevent.

A further point we would like to make is that although various theories have been used to explain research misbehavior of individual scientists, our discussion brought to light that in order for such explanations to have some minimal level of plausibility, we need to know quite a bit about the personal situation of the researcher, as well as her contextual affordances at an institutional level. The suggestion of our paper is that such knowledge is not easily obtained.

Our final point concerns the role of the Stapel case in our discussion. It should be clear that we have not tried to offer the fullest possible explanation of his fraudulent behavior. We have used Stapel merely to illustrate the kinds and amounts of facts that should be known if an explanation of research misconduct, based on any of the six theories discussed in this paper, is to have minimal plausibility.

Acknowledgements TH and RvW would like to thank the reviewers and the members of the Theoretical Philosophy reading group for their critical yet constructive comments on earlier versions of this paper.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abma, R. (2013). *De publicatiefabriek. Over de betekenis van de affaire-Stapel* (p. 183). Nijmegen: Van Tilt Uitgeverij.
- Agnew, R. (1992). Foundation for a general strain theory of crime and delinquency. *Criminology*, 30, 47–87.
- ALLEA (All European Academies). (2017). The European code of conduct for research integrity. Berlin: All European Academies.
- Anscombe, G. E. M. (2005). The causation of action. In: M. Geach & L. Gormally (Eds.), *Human life, action and ethics* (St. Andrews Studies in Philosophy and Public Affairs). Exeter: Imprint Academic.

- Bouter, L. M., Tijdink, J., Axelsen, N., Martinson, B. C., & ter Riet, G. (2016). Ranking major and minor research misbehaviors: Results from a survey among participants of four world conferences on research integrity. *Research Integrity Peer Review*, 1(17), 1–8.
- Cassam, Q. (1992). *Vices of the mind*. Oxford: Oxford University Press.
- Crain, L. A., Martinson, B. C., & Thrush, C. R. (2013). Relationships between the survey of organizational research climate (SORC) and self-reported research practices. *Science and Engineering Ethics*, 19(3), 835–850.
- Dancy, J. (2000). *Practical reality*. Oxford: Oxford University Press.
- Davis, M. S. (2003). The role of culture in research misconduct. *Accountability in Research*, 10(3), 189–201.
- De Graaf, G. (2007). Causes of corruption: Towards a contextual theory of corruption. *Public Administration Quarterly*, 31, 39–86.
- De Vries, R., Anderson, M. S., & Martinson, B. C. (2006). Normal misbehavior: Scientists talk about the ethics of research. *Journal of Empirical Research on Human Research Ethics*, 1(1), 43–50.
- Fanelli, D. (2009). How many scientists fabricate and falsify research? A systematic review and meta-analysis of survey data. *PLoS ONE*, 4(5), e5738.
- Faria, R. (2018). *Research misconduct as white-collar crime: A criminological approach*. Cham: Palgrave Macmillan.
- Fiedler, K., & Schwarz, N. (2016). Questionable research practices revisited. *Social Psychological and Personality Science*, 7(1), 45–52.
- Gilovich, T. (1991). *How we know what isn't so: The fallibility of human reason in everyday life*. New York: Free Press.
- Gunsalus, C. K. (2019). Make reports of research misconduct public. *Nature*, 570, 7.
- Hackett, E. J. (1994). A social control perspective on scientific misconduct author. *Journal of Higher Education*, 65(3), 242–260.
- Halfman, W., & Radder, H. (2015). The academic manifesto: From an occupied to a Public University. *Minerva*, 53(2), 165–187.
- Haven, T. L., Tijdink, J. K., Pasman, H. R., Widdershoven, G., Riet, G., & Bouter, L. M. (2019). Researchers' perceptions of research misbehaviours: A mixed methods study among academic researchers in Amsterdam. *Research Integrity and Peer Review*, 4(25), 1–12.
- Holtfreter, K., Reisig, M. D., Pratt, T. C., Mays, R. D. (2019) The perceived causes of research misconduct among faculty members in the natural, social, and applied sciences. *Studies in Higher Education* 1–13.
- Hren, D., Vujaklija, A., Ivanišević, R., Knežević, J., Marušić, M., & Marušić, A. (2006). Students' moral reasoning, Machiavellianism and socially desirable responding: Implications for teaching ethics and research integrity. *Medical Education*, 40(3), 269–277.
- Institute of Medicine. (2002). *Integrity in scientific research: Creating an environment that promotes responsible conduct*. Washington D.C.: National Academy of Sciences.
- Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *The American Economic Review*, 93(5), 1449–1475.
- Kahneman, D. (2011). *Thinking, fast and slow*. London: Penguin Books.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decisions under risk. *Econometrica*, 47(2), 263–291.
- Lacetera, N., & Zirulia, L. (2011). The economics of scientific misconduct. *Journal of Law Economics and Organization*, 27(3), 568–603.
- Lafollette, M. C. (2000). The evolution of the “scientific misconduct” issue: An historical overview. *Proceedings of the Society for Experimental Biology and Medicine*, 224(4), 211–215.
- Levitt Committee, Noort Committee, Drenth Committee. (2012). *Flawed science: The fraudulent research practices of social psychologist Diederik Stapel*.
- Lipton, P. (2008). *Inference to the best explanation* (2nd ed.). London and New York: Routledge.
- Lipton, P. (2009). Understanding without explanation. In H. W. de Regt, S. Leonelli, & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives*. Pittsburgh: University of Pittsburgh Press.
- Maggio, L., Dong, T., Driessen, E., & Artino, A. (2019). Factors associated with scientific misconduct and questionable research practices in health professions education. *Perspectives on Medical Education*, 8(2), 74–82.
- Martinson, B. C., Anderson, M. S., & de Vries, R. (2005). Scientists behaving badly. *Nature*, 435(7043), 737–738.
- Martinson, B. C., Crain, L. A., De Vries, R., & Anderson, M. S. (2010). The importance of organizational justice in ensuring research integrity. *Journal of Empirical Research on Human Research Ethics*, 5(3), 67–83.

- National Academies of Sciences, Engineering, and Medicine. (2017). *Fostering integrity in research*. Washington, DC: The National Academies Press. <https://doi.org/10.17226/21896>.
- Netherlands code of conduct for research integrity. (2018).
- Overman, S., Akkerman, A., & Torenvlied, R. (2016). Targets for honesty: How performance indicators shape integrity in Dutch higher education. *Public Administration*, 94(4), 1140–1154.
- Owens, D. (1989). Levels of explanation. *Mind*, 98, 57–79.
- Rajah-Kanagasabai, C. J., & Roberts, L. D. (2015). Predicting self-reported research misconduct and questionable research practices in university students using an augmented theory of planned behavior. *Frontiers in Psychology*, 6, 535.
- Schneider, B., Ehrhart, M. G., & Macey, W. H. (2013). Organizational climate and culture. *Annual Review of Psychology*, 64(1), 361–388.
- Sellin, T. (1983). *Culture and conflict*. New York: Social Science Research Council.
- Sovacool, B. K. (2008). Exploring scientific misconduct: Isolated individuals, impure institutions, or an inevitable idiom of modern science? *Journal of Bioethical Inquiry*, 5(4), 271–282.
- Stapel, D. (2014). *Faking science: A true story of academic fraud*. (N. J. L. Brown, Transl.). <https://bit.ly/3tLfKCr>.
- Steneck, N. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics*, 12(1), 53–74.
- Teelken, C. (2015). Hybridity, coping mechanisms, and academic performance management: Comparing three countries. *Public Administration* 93(2), 307–323.
- Thomas K. J., & Loughran T. A. (2014). Rational choice and prospect theory. In: G. Bruinsma & D. Weisburd (Eds.), *Encyclopedia of criminology and criminal justice*. New York: Springer.
- 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), 1–12.
- Vogel, G. (2011). Psychologist accused of fraud on astonishing scale. *Science*, 334(6056), 579.
- Wible, J. R. (1992). Fraud in science: An economic approach. *Philosophy of the Social Sciences*, 22(1), 5–27.
- Woodward, J. (2003). *Making things happen: A theory of causal explanation*. Oxford: Oxford University Press.
- Zwart, H. (2017). *Tales of research misconduct: A Lancanian diagnostics of integrity challenges in science novels* (pp. 1–254). Cham, Switzerland: Springer.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.