Chapter 9 Science and University Politics



Contents

9.1	Introduction		200
	9.1.1	Is There a Crisis in the Social Sciences?	200
	9.1.2	System Approach	201
9.2	Replication Crisis		202
	9.2.1	What Is a 'Replication Crisis'?	202
	9.2.2	Lack of Replication Studies	202
	9.2.3	Weak Research Rigor	205
	9.2.4	Failing Replication?	205
9.3	Publication Pressure		207
	9.3.1	Value for Money	207
	9.3.2	Publish or Perish	208
	9.3.3	Intrinsic Versus Extrinsic Motivation	210
9.4	Perverse Incentives		210
	9.4.1	Reward Systems	210
	9.4.2	Matthew Effect	211
	9.4.3	Gap Between Tenured and Contingent Faculty Members	212
	9.4.4	Gender Gap?	213
	9.4.5	Summing Up	214
9.5	Teaching Research Ethics and Integrity		214
	9.5.1	Ethics and Integrity Education	214
	9.5.2	Reactive Education	215
	9.5.3	Proactive Education	216
	9.5.4	Reflexive Education	216
	9.5.5	Should Misconduct Be Criminalized?	216
	9.5.6	In Sum	217
9.6	Conclusions		217
	9.6.1	Summary	217
	9.6.2	Discussion	218

Electronic Supplementary Material: The online version of this chapter (https://doi.org/10.1007/978-3-030-48415-6_9) contains supplementary material, which is available to authorized users.

Case Study: How to Start a Fire	218
Assignment.	220
Suggested Reading	220
References	221

After Reading This Chapter, You Will:

- · Understand how political factors impact modern science
- Appreciate in what ways the replication crisis endangers the values of science
- Know how publication pressure and perverse incentives challenge scientific practices
- · See why teaching ethics requires reactive, proactive, and reflexive education

Keywords Adjunctification of higher education · Competition for prestige · Degree of freedom · False positives · h-index · Impact factor · Matthew-effect · Motivation (intrinsic v/s extrinsic) · New public management · Nomothetic · Paradigm · Performance indicators · Performance-based research funding system (prfs) · Perverse incentives · Publication bias · Publication pressure · Research rigor · Replication (exact v/s conceptual) · Replication crisis · Salami slicing technique · System theory

9.1 Introduction

9.1.1 Is There a Crisis in the Social Sciences?

The Guardian, a leading British newspaper, had a long running series on abuses and exploitation in academia titled 'Academic anonymous'. From this series, we draw a few examples. In one piece (30 June 2017), the work of an unnamed researcher is described. It shows how the researcher tailored their papers not to fit the data, but to make the papers tailor fit for publication in prestigious journals. Another researcher left out any data interpretations that could raise questions with the journal editors. Their supervisor told them it could lead reviewers to turn it down. In yet another example, a research supervisor informed their students that research begins not with asking questions but with selecting suitable, high-ranking journals and defining subjects that might fit into those journals. All these examples illustrate the less than perfect publication behavior currently practiced in some corners of academia.

It was precisely experience with such practices that incited cognitive neuroscientist Chris Chambers to write *The Seven Deadly Sins of Psychology* (2017), an indictment of the numerous research misbehaviors found in the field of psychology, many of which already have been discussed in previous chapters of this book (see especially Chaps. 6 and 7), ranging from forms of bias and unreliability to fraud or even corruption. After working for nearly 10 years in the field, Chambers in an op-ed for *The Guardian* (May 9th 2017), said that he understood that a psychologist's mission wasn't about truth seeking, it was about 'crunching through as many experiments as possible as quickly as possible, finding ways to make ambiguous data look beautiful, publishing frequently in prestigious journals [...] winning large public grants, and basically getting as famous and powerful as possible'.

Whether we will call the state of affair psychology finds itself in a 'crisis,' as some do, or more of a 'challenge,' may be a matter of opinion. Chambers found that not enough had been done to address the forces that fueled the problem (see our discussion of 'academic capitalism' in Chap. 8). An appeal to research ethics could be instrumental in transforming the field psychology in helping it to overcome this (perceived) crisis.

Whether this is the case will be the subject of debate in this chapter. We realize, however, that many of the discussions in this chapter are not part of a student's lived experience. Many of the themes addressed here may seem 'abstract,' far removed from the reality of everyday life. To a degree this is true, but to a degree it isn't, because the pressures and incentives, and political considerations that make up the fabric of modern academic life will reveal themselves in the educational practices in which students are immersed. This is why we have chosen to dedicate one chapter of our book to some of these broader topics.

9.1.2 System Approach

In this chapter, we will broaden our perspective on research ethics, and move away from considering how or why individual researchers decide to break from established norms and values, or from examining what happens when rigid guidelines meant to steer our ethical behaviors are not followed. Instead, we will look at the interrelated and interdependent parts that make up the whole of the scientific enterprise, of which ethics is just one element. This holistic approach is often referred to as a *systems perspective* (see Parson 1951; Wiener 1965; Maturena and Varela 1980). By zooming out to a lens that can capture the system as a whole, we can consider how scientific practices and ethics mutually influence one another. This allows us to explore some of the most important internal factors that regulate scientific operations and shape our understanding of how ethics play a role.

We begin with an exploration of what is currently considered one of the more challenging problems in the social sciences, namely the *replication crisis*. Next, we will research two factors that are often closely connected to said crisis, both believed to have a corroding influence on ethics: *publication pressure* and *perverse incentives*. In the final section, we will focus on the role of teaching research ethics in the future as a means of countering the corroding effects of misconduct.

9.2 Replication Crisis

9.2.1 What Is a 'Replication Crisis'?

The term *replication crisis* came into circulation around 2010, when it was observed that often, when studies were reproduced, the same findings could not be replicated. The replicated studies would find a smaller or larger effect than originally claimed, or none at all, or the direction of the effect had changed, or perhaps an entirely different effect was found altogether.

Derived from the natural sciences, the requirement that credible knowledge must consist of reproducible findings implies that studies carried out under *the same conditions should generate the same results*. Findings that cannot be reproduced are considered *chance findings, or false positives*, and do not belong in the scientific body of knowledge (Makel 2017).

The requirement of replicability is based on a *nomethetic* approach, meaning that it should be possible to discover laws that explain objective phenomena. Although not every social scientist accepts the requirement of replicability and its underlying nomothetic notion, those who do are committed to experimentation, clinical trials, and hypothesis testing under controlled conditions. The observation of a replication crisis calls into question the credibility of such research findings.

In fact, many of Daniel Kahneman's well known social priming studies (as detailed in *Thinking, Fast and Slow*, 2011) failed in replication, likely because the sample sizes had been too small. 'I placed too much faith in underpowered studies,' Nobel prize laurate Kahneman famously admitted (*Retraction Watch*, February 20th 2017). Kahnemann is but one of many whose studies failed to be replicated (see Hughes 2018, for more examples).

What constitutes the 'replication crisis' is a digression from the ideal of reproducibility, resulting in the view that many research findings in certain disciplines (notably psychology, but also medicine) are unreliable, or perhaps even false. To better grasp the issues at hand, we discuss three aspects of the replication crisis: (1) lack of replication studies, (2) weak research rigor, and (3) failing replication.

9.2.2 Lack of Replication Studies

While replicability is accepted as a 'fundamental principle' by a large portion of the social science community, very few replication studies are actually carried out. This was noticed by Smith as early as 1970, while more recently, Makel et al. (2012) found an overall replication rate of psychological research at just over 1%. Additionally, many of these replication studies were performed by the original author. Why is there so little interest in replication studies? Let's look into two possible reasons.

One rationale is *theoretical sectarianism* (favoritism of one's own theories, and prejudice or discrimination against rival ones), to which many social scientific disciplines are prone. Theoretical sectarianism occurs because researchers operate under specific theoretical assumptions, or *paradigms*, as they were defined by Thomas Kuhn (1970). Kuhn conceptualized a change in theoretical assumptions as a *paradigm shift*. The problem being, paradigm shifts aren't often universal, with different disciplines regularly operating under different paradigms. Furthermore, there is often little common ground to be found between differing paradigms.

Without common ground (shared theories, mutually referenced authors, agreed upon methodologies), researchers are more inclined to confirm findings in their own domain than to disconfirm those of their peers in other domains. This implies that the use of unfamiliar methods make an interdisciplinary researcher performing a replication study to be more vulnerable to the possibility (or accusation) of misunderstanding (see Box 9.1 for an example thereof).

Box 9.1: Irreproducible Tears?

The issue of 14 January 2011 of the journal *Science* ran a paper by a team of Israeli scholars from the Weizmann Institute of Science at Rehovot, that claimed that human tears my serve a chemosignaling function. After a team of Dutch scholars subsequently failed in their attempt to replicate the experiment described in the original article, a dispute broke out between the two groups of scientists, each accusing the other of misrepresentation. The case, which we describe below in some detail, illustrates the difficulties of replication in the social sciences.

The Israeli team, headed by Naom Sobel, hypothesized that human tears serve a *chemosignaling function*, such that men smelling or breathing in female tears become sexually less aroused. The idea is that tears contain certain chemicals that may influence brain activity in (heterosexual) men. In order to test this hypothesis, the researchers collected 'donor tears' from women who had watched sad films. They then exposed a group of males in a *within-subjects design* to these tears as well as to a substitute substance. After exposure, the participants were asked to rate their sexual attraction to pictures of female faces. In a second experiment, the researchers added another dependent variable, namely levels of psychophysiological arousal. The various studies revealed that exposure to female tears indeed has an effect, both on self-reported levels of sexual arousal (modest) and on objective psychophysiological expression (more pronounced) (Gelstein et al. 2011).

A group of mainly Dutch psychologists attempted to replicate these findings. One experiment was set up as an exact replication, the other two had

Box 9.1 (continued)

alterations. Thus, in one condition, male participants were not only asked to rate sexual attraction, but also whether or not they would be willing to date the females in the pictures. In a second series of experiments, the researchers changed the design from a within-subjects study (which had low power), to a larger, *combined within and between subject design*. However, they now asked the subjects to not only rate the pictures of a female face, but also of the whole body as a measure of 'arousal.'

Their attempts to 'replicate and extend' the original studies failed. They found no support for the chemosignaling function hypothesis, which they considered a possible *false finding* (Gračanin et al. 2017a, b, p. 149). The authors proposed instead that tears are functional in a *social* context, and that crying is a 'self-soothing strategy' (Gračanin et al. 2014).

Principal investigator Naom Sobel (2017) of the original Israeli study responded to the failed attempt at replication, arguing that the replication studies were not really replications of the original at all. The researchers had not operated from a proper chemosignaling laboratory, he argued. Further, he stated they had used different test materials (other films), which communicated a different feeling (not sadness), and had used combined datasets in an 'inappropriate manner.' Had they used the 'appropriate techniques,' Sobel believed they would have found that the data the Dutch team collected actually supported the original hypothesis. This prompted a further response from the Dutch replication team (Gračanin et al. 2017a, b), who argued that a theory that only holds under very specific circumstances is likely not a very good theory.

This dispute is of interest because it leaves open any of the following alternatives, namely that this 'failure to replicate' represents:

- 1. An instance of theoretical sectarianism
- 2. The social sciences' context dependency
- 3. Weak research rigor

Which of these alternatives do you believe is most likely?

A second rationale may be that journals prefer to publish 'newsworthy findings' over what is 'already known' (see the section on dissemination bias in Chap. 6 for further discussion). For this same reason, it is more difficult to get funding for replication studies, especially *exact replication studies* (same methodology, same conditions), as opposed to *conceptual replication studies* (same conceptualization, different methodology or different conditions).

The net result is that in the absence of adequate replication studies, we are much less sure that positive findings are in fact positive, and not accidental (see Fanelli 2010a for a discussion).

9.2.3 Weak Research Rigor

The replication crisis may be further worsened by *weak research rigor*. In 2011, Simmons et al. published a now famous study that shows how easy it is to produce false positives if the researcher's *degree of freedom* is manipulated. Degrees of freedom pertain to a series of methodological choices researchers must make over the course of collecting and analyzing data, including the selection of dependent variables, determining sample size, using covariates (independent variables), and reporting subsets of experimental conditions (Simmons et al. 2011).

While staying within the accepted boundary of a .05 p-value, the authors succeeded in 'proving,' in a real experimental study with real subjects, the unlikely conclusion that listening to certain types of music makes you feel younger, and the obviously false conclusion that listening to certain other types of music *actually* makes you younger. They arrived at these conclusions by tacitly manipulating the researcher's degrees of freedom (such as changing the number of participants without reporting this, or reporting on certain measures only). A reviewer unaware of these manipulations would have to accept the 'age effect' (that you actually become younger by listening to certain music) as genuine.

The point the authors sought to make was that as long as reviewers and readers are not informed about the researchers' choices within their degrees of freedom (which can in and by themselves be legitimate), they cannot reasonably separate false from true findings. They therefore recommend more transparency about the choices made by researchers to avoid the creation of false positives.

9.2.4 Failing Replication?

Another seemingly significant blow was delivered in 2015, when a number of researchers undertook a large scale attempt to replicate a swath of psychological studies. The result were sobering. At the insistence of Brian Nosek, a group of 270 authors (worldwide) replicated one hundred psychological studies, published in the course of 1 year (2008), in three top ranking journals (Open Science Collaboration 2015).

Using a uniform replication protocol, all studies were carried out as *exact replications* as much as possible. The results were collected, assessed, and independently reviewed. Comparing the original studies with the replication studies, the researchers looked at significance and p-values, effect size, subjective assessments, and meta-analyses of effect size. What they found was that collectively, these indicators revealed that the replications produced significantly weaker evidence than the original findings. For example, the mean size effect of the replication studies were roughly half the magnitude of original the mean size effect (see Fig. 9.1).

The Open Science Collaboration project caused shockwaves that resonated beyond the scientific community. Newspapers all over the world reported that psychological studies fail to replicate. The authors themselves were much more careful,



Fig. 9.1 Original study effect size versus replication effect size. (© Science, 28 Aug. 2015, issue 6251)

however. They argued that their study neither proved nor disproved anything. 'The original studies examined here offered tentative evidence; the replications we conducted offered additional, confirmatory evidence. In some cases, the replications increase confidence in the reliability of the original results; in other cases, the replications suggest that more investigation is needed to establish the validity of the original findings. Scientific progress is a cumulative process of uncertainty reduction that can only succeed if science itself remains the greatest skeptic of its explanatory claims' (Open Science Collaboration 2015, p. 4716–7).

Maxwell et al. (2015) similarly argued that when replication studies fail to show significant results, this should not lead to the premature conclusion that the original study was somehow faulty or flawed. In a reanalysis of the data collected by the Open Science team, Van Bavel et al. (2016) were able to demonstrate that some of the failure to reproduce the same findings could be attributed to *contextual factors*, having to do with social and cultural differences between the countries where the studies were carried out, illustrating just how difficult it is to replicate social science studies.

9.3 Publication Pressure

9.3.1 Value for Money

In Chap. 8, we discussed how in the last quarter of the twentieth century, globalization and neoliberalism set in motion a trend towards 'valorization' of science, that emphasizes the monetary value of scientific knowledge. Ties with external parties (both commercial organizations and governmental bodies) strengthened, from which a variety of conflicts emerged. But 'valorization' also impacted the way science was organized internally, including publication behavior, to which we now turn our attention.

The wish to demonstrate to the tax paying public that scientists produce 'value for money' found its way into an administrative logic that demands adherence to uniform, quantifiable norms. University administrators and policy makers began thinking about ways to measure a researcher's output in terms of 'objective numbers'. One objective criterion they soon came up with was a count of articles published by individual researchers in prestigious journals. Later this was extended to a count of citations received. Both could be used as an approximation of a researcher's objective impact on the scientific community.

These measures resulted in the establishment of new standards, such as a journal's *impact factor*, a researcher's *citation impact*, and their so-called *h-index* (see Box 9.2). Any of these standards, with a special focus on publishing success, became a measuring stick for success in academia (Barnard-Brak et al. 2011). Greater output and more citations equaled greater academic value. Unwittingly, these norms transformed the tenured job market in academia into a 'market for prestige' (Garvin

Box 9.2: H-index (Adapted from Wikipedia)

The *h-index* (named after its inventor, Jorge Hirsch) is an author-level metric that attempts to measure both the productivity and citation impact of a scientist's publications. The index is based on the scientist's most cited papers and the number of times those papers were cited in other publications. The index can also be applied to the productivity and impact of a scholarly journal, as well as a group of scientists residing within the same department, university, or even country. The values on the h-index vary greatly from discipline to discipline, where the number of scientists, published papers, and citations strongly differ. In physics, for example, the average h-value for a fulltime professor is between 15 and 20, in economics its 7.6, in sociology its 3.7. In many academic communities, it was long believed that better scholarship equaled higher h-factors, but recent discussion flared up about whether h-values actually represent quality.



Fig. 9.2 Publication pressure

1980), thereby creating of whole series of new questions, which we will outline below (Fig. 9.2).

9.3.2 Publish or Perish

Has the pressure to publish in 'high impact journals' altered research norms? Moustafa (2014, p. 139) observed that the impact factor became 'a major detrimental factor of quality, creating huge pressures on authors, editors, stakeholders and funders.' More tragically, Moustafa notes, is that impact factor has also become the condition for allocating government funding to entire institutions in some countries.

Siegel and Baveye (2010) reason that scholars who wish to meet publication expectations will resort to a variety of techniques to increase their output and crank up their citation ranking. These include the use of *co-authorships* and so called *gift authorships* (author does not contribute to the research, or not significantly, but is included out of courtesy and in the expectation of reciprocity). Additionally, *salami slicing* techniques (slicing research such that several different papers can be written, all slightly varying around the same theme) and *extensive referencing* ('I cite-you, you-cite-me') are employed to meet publication expectations (Box 9.3).

Fanelli and Larivière (2016) researched why publication rates have increased in the last quarter of the twentieth century. Looking into the work of over 40,000 researches in Western countries, whose profiles were drawn from the Web of Science, they compared the publishing frequency between 1900 and 1998. They found that the average number of papers published throughout the twentieth century remained stable for most disciplines, and then visibly increased after 1980. However,

Box 9.3: Slicing and Dicing: A Dilemma

A well-respected colleague proudly explains to you that he has managed to produce 12 publications out of the one dataset he collected for his dissertation. This is a particularly interesting achievement, as it involves a dataset with only 232 respondents to a four-page survey.

How do you respond?

- 1. I think this is a great example to follow and I ask him how he achieved it.
- 2. I cannot imagine each of these 12 papers having a unique contribution and vow never to go down this route.
- 3. I tell the colleague that this is bad science and that I strongly disapprove of their actions.
- 4. I think this is bad practice and is tainting the reputation of science. I inform the editors of at least the most recent of the 12 publications.

(Adapted from the Erasmus Dilemma Game)

so did the number of co-authors. *Fractional productivity* (the productivity of one researcher's work spread out over multiple co-authored papers) remained stable. From this, the authors concluded that the widespread belief that pressure to publish causes the scientific literature to be flooded with salami-sliced, trivial, and false results is in fact incorrect, or at least exaggerated.

In a similar examination, van Wesel (2016) looked at output of journals rather than individual researchers. He examined 50 high impact scientific publication outlets selected from a variety of disciplines (including medicine, physiology, and psychology), to see whether publication behavior changed between 1997 and 2012. Van Wesel compared the number of authors listed, the amount of references included, but also the text and abstract lengths, and even the presence of a colon in the title (as indicators of a paper's 'citability'). He found similar patterns as Fanelli and Larivière (2016), including a growing number of co-authors and increased referencing. Aware that these patterns could also be attributed to a change in editorial policies, van Wesel (2016, p. 212-3) believes it is nevertheless 'not unrealistic to link the observed changes in publication behavior to a change in evaluation criteria.'

While the studies discussed above do not provide evidence that 'pressure to publish' leads to fraudulent behavior (Fanelli 2010b), there is reason to believe that publication norms have changed. Authors seem much more strategic in their publication behavior. Due to an awareness of the necessity to publish in order to further their academic careers, it seems likely that researchers plan their publications to meet these goals.

9.3.3 Intrinsic Versus Extrinsic Motivation

Does increasingly 'strategic' and 'planned' publishing behavior means that authors are less intrinsically motivated? Hangel and Schmidt-Pfister (2017) interviewed over 90 researchers in Germany who were in different stages of their careers (PhD students, post-doc, tenured professors). The authors asked their participants frankly: 'Why do you publish?' They found that the most common reason to publish contained both internal and external motivations, namely: (1) to communicate interesting research results; (2) to gain recognition among peers; (3) because they enjoy writing, and (4) to obtain funding in order to secure future research.

All researchers draw from a mix of these motivations, albeit differently in different stages of their careers. Those still in the early stages of their careers, when they are more dependent on their supervisors, are often highly aware of the incentive that one *has to* publish for mere survival. They cite motives 2 and 4 as most important. In the next stage of their career, many still feel the pressure to publish, but now more as a means to an end. These researchers seem more capable of enjoying the research process itself and more often cite reasons 1 and 3. Tenured professors almost always cite reason 4, claiming to publish for educational purposes and for academic survival.

9.4 Perverse Incentives

9.4.1 Reward Systems

Publication pressure represents one of the factors that contribute to what some believe to be a precarious situation in the social sciences. Ambition, external pressure, and weak research rigor seem to highlight an element of 'crisis' mentioned earlier in this chapter. Outside of publication pressure, the other factor consists of reward systems that aim to stimulate academic quality and simultaneously deal with decreased public funding.

Acting on a desire to transform universities along neoliberal lines into efficient, productive, and outstanding institutions, *new public management* administrators (see Box 9.4) set up national reward systems to select the best researches. *Performance-based Research Funding Systems* (PRFS) have since been used to evaluate the quality of research proposals. They are based on the rationale that 'funding should flow to the institutions where performance is manifest' (Herbst 2007, p. 90).

These systems take into account any number of 'performance indicators,' such as the output of a research group, their citations received, their international ranking, and the judgements of their peers. For individual researchers, the number of successful grant applications is counted, along with their level of participation in international research associations, the number of keynote addresses they've made, board memberships, awards bestowed, and even their perceived societal impact (as expressed, for example, in their role as advisors for social organizations).

Box 9.4: New Public Management

New Public Management (NPM) is the late twentieth century approach to public service organizations that suggested they be run like businesses. It is based on principles of expanding managerial freedom, flexibility of organizational structures, shifting staff and job conditions, emphasis on output and decrease of input ('cost-effective management'), and increase of efficiency (see Christopher Hood, 'A Public Management for All Seasons?', 1991).

Perhaps unsurprisingly, researchers faced with the prospect of losing funding started to seek out ways to increase their performance accordingly. In an attempt to keep pace, or even outdo their colleagues, researchers began spending more time writing research proposals, but since each of these have a limited chance of being honored, the 'battle for efficiency' effectively resulted in an even greater waste of public resources.

PRFSs and other performance-based evaluation systems have unintended, so called *perverse effects*, when they encourage behavior that runs contrary to the original intentions (stimulation of excellence resulting in inefficient use of public funding) or invigorate undesirable behavior (competition for prestige, see Hicks 2012). Bouter (2015, p. 157) paints a grim picture of how certain incentives lead to a 'monoculture focused on citation scores, short term economic gains, and government-defined growth sectors,' with young talented researchers not being scouted. In the following sections, we discuss areas where such incentives can have perverse effects (see also Stone 2002).

9.4.2 Matthew Effect

Sociologist Robert Merton (1968) reasoned how symbolic, as well as material rewards in science will have an accumulative effect. If a researcher 'scores' on any one of the criteria mentioned above, their chances of getting funded improve, and this effect will accumulate over time (also in a negative sense: if you don't score, your chances diminish accordingly). The result is that 'eminent scientists get disproportionately great credit for their contributions to science while relatively unknown ones tend to get disproportionately little for their occasionally comparable contributions' (Merton 1988, p. 607). This is called the *Matthew Effect*, named after the Gospel of Matthew 25:29: 'For to every one who has will more be given, and he will have abundance; but from him who has not, even what he has will be taken away.'

There is some evidence that indicators of past performance correlate positively with chances of getting a grant application funded. This has been found in the fields of economics and the social sciences in the Netherlands (van den Besselaar and Leydesdorff 2009; Bornmann et al. 2010), although the correlation is low and is not consistent across all disciplines.

More worryingly, since the Matthew Effect contains an element of self-fulfilling prophesy in it, there is a danger of misrecognition attached to it, especially with young academics. When academics are overly assessed in the early stages of their careers, precocious students will have a far better chance of surviving the competitive struggle over their late blooming peers, who may be just as brilliant.

9.4.3 Gap Between Tenured and Contingent Faculty Members

A further concern is that certain policies will negatively affect the academic labor market (Schwartz 2014). There are strong indications that neoliberal university politics, though meant to reward academic excellence, contribute to a divide between tenured staff with permanent positions and contingent faculty members who work on temporary contracts. This process is referred to as the *adjunctification of higher education*, or academia's overreliance on temporary, non-tenured faculty members (Curnalia and Mermer 2018).

There are a number of reasons for this *adjunctification*. Increased focus on retention, career outcomes, and resource acquisition brought about a reduction in tenured positions. Once regarded as essential to protect academic freedom in the pursuit of knowledge, today tenured positions making up sometimes as little as 20 to 25% of all faculty staff in Western universities. The majority of instructors and teachers are hired on a contingent basis. These have little prospect of getting tenured but are assessed by and large along the same criteria as tenured staff, while their possibilities for doing research (and getting promoted) are diminishing. Teaching and research are thus increasingly being undertaken by different kinds of faculty (Finkelstein 2014). If nothing changes in the near future, it is to be expected that this divide will only grow deeper (see Dobbie and Robinson 2008) (Fig. 9.3).

In a survey among some 1500 higher education professors, deans, governing board members, campus administrators, policymakers, and other stakeholders in the United States, Kezar et al. (2015) found general agreement that the present system is untenable. It threatens academic freedom and undervalues teaching through its disproportionate emphasis on research. Most respondents agreed on the necessity to restructure teaching positions. More full-time faculty, differentiation of responsibilities, and an overarching need to restore professionalism to the role of faculty were among the most pressing urges uttered by administrators and professors alike.

Below, we give the floor to two adjunct faculty who experienced this gap between tenured and contingent up close. One, a Dutch scholar in the social sciences, left academia after years of working on a temporary basis. In an email to her colleagues (quoted here with permission), she wrote: 'I experience a large disconnect between what you are paid to do as a temporary staffer (teaching) and what determines your career options (research). This disconnect means that many contingent faculty members are almost forced to put in a lot of additional hours (i.e., evenings,



Fig. 9.3 If nothing changes, the gap will grow

weekends, using holidays to write papers, taking up parental/care leave to decrease the teaching load and use that time to work). If you do not want to go along with this, your research output will stay 'behind' compared to the competition, which makes you less attractive and decreases your chances on permanent positions, promotions, grants, etc.'

Similarly, an adjunct faculty member in England was quoted as saying: 'I've watched brilliant friends be employed for two or three consecutive years with demanding teaching loads, travelling to cities hundreds of miles away with sharing childcare, only to be dropped for someone else with a more illustrious publication record' (*The Guardian*, July 14th 2017).

9.4.4 Gender Gap?

Historically, women have been underrepresented in all disciplines in academia and at all stages in their careers. Many industrialized countries have adopted strong gender equality programs in research and innovation, and the gender gap has since grown smaller (Ceci et al. 2014). Despite this, there is still a 'pipeline leakage,' meaning that on the way to the top, women drop out more often than men (Huyer 2015). From the 2015 report *She Figures*, issued by the European Commission, we learn that in 2011 some 50% of all students in the social sciences and law in the EU, about 30–40% of researchers in these fields, and 29% 'top level' researchers therein are women (with considerable differences per country).

Is the demise of the gender gap in higher education (in industrialized countries) an effect of government policies or does it happen independently? Ginther and Kahn (2009) compared the careers of male and female researchers in the US

between 1973 and 2001. They examined the probability of obtaining a tenure track after getting a PhD and found that women were still less likely to take tenure positions in science. This was explained by 'fertility decisions' rather than reward systems – meaning that 'women must face a choice between having children or succeeding in their scientific careers, while men do not face the same choices' (Ginther and Kahn 2009, p. 183).

In non-industrialized countries (for example in parts of Africa), the gender gap persists, and it may also not be ascribed to government policies but rather cultural expectations. Ogbogu (2011), reviewing the gender gap in higher education in Nigeria, notes that recruitment and selection practices in their universities do not discriminate against women. Instead, factors such as lack of mentoring, poor compensation, family responsibilities, 'and the ideology that women should have low career aspirations' accounted for the observed disparity in academia.

9.4.5 Summing Up

In an attempt to keep up with and respond to worldwide changes in the political and economic landscapes within which universities operate, university administrators and government policy makers have set up procedures that aim to enhance productivity and excellence. These policies have had (and continue to have) intended as well as unintended consequences, resulting in a number of fundamental changes in the ways universities conduct research and provide education. These changes themselves pose new questions and challenges that must be addressed (Fig. 9.4).

9.5 Teaching Research Ethics and Integrity

9.5.1 Ethics and Integrity Education

There is little doubt that ethics and integrity education is becoming increasingly important in universities, not least because of increasing demands and competition in the academic field (Brall et al. 2017). Universities have a commitment to prepare, guide, and mentor students through a litany of ethical issues; combatting scientific misconduct, addressing questionable research practices, applying specific procedures and regulations, learning to deal with newfound responsibilities belonging to certain roles, knowing how to accommodate a diversity of perspectives, and learning how to deal with external pressure (Naimi 2007). We concur with Resnik (1998, p. 174) that the question is not whether 'ethics be taught?', but 'how can ethics be taught?' In the sections below, we briefly discuss three broad approaches for how ethics can be taught.

Fig. 9.4 © Cartoon by John Stuart Clark



9.5.2 Reactive Education

This approach is focused on the prevention of misconduct and misuse of procedures. To accomplish *reactive education*, many research institutions offer *casebased approaches* in their curricula (with either real or hypothetical dilemmas), where students learn to make judgement calls through structured discussions (Sponholz 2000). These discussions allow students to 'evaluate conventions, define responsibilities, articulate positions on different issues, and acquire some facility at using a framework for ethical decision making' (Stern and Elliott 1997). Canary (2007) shows that these approaches are successful in enhancing moral sensitivity, moral judgements, moral motivation, and moral character in students.

One such development consists of the development of eLearning tools. In a discussion of the issues related to the emergence of computer-based learning, Esposito (2012) finds that open networked learning environments 'encourage a participatory research approach and therefore foster creative suggestions and shared solutions from participants in an evolving landscape of ethical opportunities and challenges' (p.323).

9.5.3 Proactive Education

This approach is focused on preparing students to actively participate in complex research environments and providing skills for adapting to changes in research policies. *Proactive education* makes use of role playing and simulation settings, which are used to train students to contribute to research themselves.

Sweet (1999) and Karkowski (2010) discuss the function of simulated review procedures with mock research proposals, used to prepare students to produce higher-quality research proposals themselves, with both authors finding these procedures helpful. Löfström (2016) discusses the use of role-playing strategies in academic integrity education, specifically staging panel discussions of realistic cases, which act as added value in facilitating perspective-taking and the broadening of a student's worldview.

9.5.4 Reflexive Education

This approach aims at developing ethical awareness not restricted to institutional procedures. Rather working within a broader definition of research ethics that includes social, political, and moral dimensions, Von Unger (2016) emphasizes the need for more critical dialogue in ethics education. As an example, a course format is discussed where sociology students were trained to reflect on a case that had political relevance. The students collected their own data, engaged in critical inquiry, learned to formulate and revise their own assumptions, and thus learned to become more self-critical, a cornerstone of ethical decision-making.

9.5.5 Should Misconduct Be Criminalized?

A final note on the question of what to do when teaching ethics and integrity fails to achieve its goals. Should misconduct be criminalized? Until very recently, misconduct rarely led to litigation. Even Diederick Stapel, whose case was discussed in Chap. 5, was never brought before a court, though he did lose his job and a large number of his articles were retracted.

A critical issue in deciding whether research misconduct should be subject to criminal law is how it is defined, argue Dal-Ré et al. (2020). Should it only cover well-known forms of fraud, such as plagiarism, fabrication, and falsifying, or should it extend to questionable research practices, such as selective reporting? This question is important, because while criminalization could deter everything that is regarded as research misconduct, it could simultaneously lead to normalization of what is *not* considered misconduct.

Dal-Ré et al. (2020, p. 9) admit that a research integrity organization with global authority will not emerge any time soon, but they are hopeful that 'a strong statement that is widely supported can unify and inspire the field.'

9.5.6 In Sum

All of the approaches discussed above contribute to promoting a stronger understanding of research ethics and integrity in students. We do not want to suggest that one of these is better than another. We merely want to argue that they all fulfill their own role in ethics and integrity education, and that universities and educators alike have a never-ending obligation to prepare students the best they can, so that they can prepare the next generation the best *they* can.

9.6 Conclusions

9.6.1 Summary

In this chapter, we examined research ethics in the social sciences from the perspective of a *systems approach*. We situated universities in the dynamic interplay of political and economic forces, and, more specifically, we discussed the influence of new public management politics on university policies. This led us to probe whether there truly is a research misconduct 'crisis' going on in the social sciences.

First, we investigated whether the social sciences suffered from a *replication crisis* – the problem that the few studies that have even been reproduced often fail to be replicated. Along those lines, we also observed that many studies suffer from *weak research rigor* where true findings cannot be distinguished from *false positives*.

Second, we discussed the impact of *perverse incentives* on scientific practices. We found that contrary to what is often suggested, these incentives may not incite fraudulent behavior directly. But there is evidence that they link to several other undesirable trends, including a trend towards *adjunctification of higher education*, or increased prevalence of educators on temporary assignments, which has an indirect impact on research ethics.

Finally, we discussed three different approaches for teaching research ethics and integrity in universities. These approaches can be used to help students come to grips with ethical questions from a *reactive, proactive,* and *reflexive* point of view.

9.6.2 Discussion

This chapter addresses some of the more fundamental problems in the social sciences, specifically those relating to the political and economic forces at play. These forces have resulted in a 'crisis' of sorts, and a drastic restructuring of universities. However, we do not offer political or economic solutions to these issues, which would fall outside the scope of this book.

The second question we did not address is whether teaching ethics and integrity can help solve some of these problems, or could possibly even contribute to them. Some argue that research ethics in the social sciences needs to be modelled after similar, standing practices in the medical sciences, and that governing bodies, such as IRBs, will have to play a more prominent role in upholding professional standards. Others resist this idea based on the objection that a highly professionalized scientific enterprise undermines scientific freedom and creativity (see Resnik 1998, p. 177). There are others still that argue that the formalization of ethical procedures achieves the opposite of what they aim to achieve. Increasingly formalized research ethics structures cause a rupture in the relationship between 'following rules' and 'acting ethically,' and the result of which is called *ethics creep* (Haggerty 2004).

Let us conclude by asking what are your thoughts on this question. Is your institution doing enough, or perhaps even too much, with regard to research ethics? Do you feel prepared to tackle the questions discussed in this book?

Case Study: How to Start a Fire

'Education is not the filling of a pail, but the lighting of a fire'.

In a March 24th, 2017 article in *Psychology Today*, Jesse Marczyk explained how he tried to motivate his students to always identify how they can improve. To achieve this end, he reported to using 'a unique kind of assessment policy,' allowing revisions after grades were received (Marczyk 2017).

From an educational perspective, this seemed 'the reasonable thing to do,' but from a professional perspective, it's plain 'stupid.' When declaring why, Marczyk noted he spent far too much time reading, commenting, and grading papers. For every 100 students, he needed 8 to 16 hours per test, time he argued should be better spent writing grant applications. That is, if he wanted to apply for a tenured track, because hiring committees 'aren't all that concerned with my students' learning outcomes.' Marczyk's article drives home the point that the commodification and underappreciation of education is directly linked to sciences' replication crisis, and that perverse incentives incite academics to behave in irresponsible ways. When teachers are evaluated on the number of students that pass their tests, they will start 'teaching to the test.' 'Rather than being taught, say, chemistry, *students begin to get taught how to take a chemistry test*, and the two are decidedly not the same thing.'

In this case, we shall look at the consequences of employing a 'technical' perspective on teaching in academic education that emphasizes its instrumental use only (teaching should serve well defined and measurable goals as efficiently as possible). We briefly discuss three authors who criticized, each in their own ways, the assumptions underlying such a perspective.

We begin with Gert Biesta (2007), who offers a series of philosophical arguments against the idea that educational practices should be modeled after the evidence-based approaches found in the field of medicine. Some educationalists, Biesta write, believe that much like the practices in medicine, teachers should rely on evidence from research about what 'works' in education, and establish similar 'evidence-based educationalist practices.' Biesta disagrees with these educationalists.

Underlying their 'evidence-based philosophy' are several untenable assumptions, Biesta argues. First, that there exists 'neutral frameworks' that allow for the assessment of outcomes (namely what is 'effective' in education). In reality, what 'works' in education is not something someone can establish beforehand objectively, it is something that is constructed in the process by teachers and students together.

In a similar vein, James Kennedy (2016), who adheres to Biesta's critique, argues against the idea that effective learning at universities can be both measured and guaranteed. On the contrary, says Kennedy. Learning is not 'a given,' nor is teaching a science – it is an art (2016, p. 35).

But there is room for improvement at universities, Kennedy acknowledges. One suggested area relates to the focus in recent years toward articulating the social impact of research. Universities should strive to better articulate the social impact of teaching and embed it into academic curricula. This would entail providing more room for student engagement, linking their inner concerns and passions to the needs of the world. Kennedy advocates for a new emphasis on student motivation, and a shift towards learner-led programs, that allow students more control over their learning processes.

Finally, we discuss a study by Andrew Boocock (2013), who interviewed administrators, departments heads, and lecturers at a UK college about their experiences with 'performance indicators' (statistics which compare universities and colleges against benchmarks for a number of preestablished, politically motivated outcomes). In this case, the desired outcomes were as follows: (a) better retention rates of students in underrepresented groups, and (b) a higher percentage of students continuing higher education. Certain measures were taken to achieve these goals. How did these indicators impact the motivation of the actors involved? Administrators often saw great advantages in striving toward performance indicator benchmarks. Student retention rates improved, their achievements went up, guidance and support systems for students were used more efficiently, as were their learning plans.

Teachers perceived a different reality, however. They noticed several side effects of focusing on performance indicators. Student attendance rates and punctuality worsened, which they attributed to the introduction of inertial, bureaucratic disciplinary procedures (Boocock 2013, p. 313). Teachers also experienced a perception of powerlessness and loss of authority, with one teacher describing it as the 'de-professionalized status of lecturers.' Their experience was troubling indeed. 'You're also under pressure in that if you are underperforming they can close your course or put it on special measures internally ... We are very much ruled by numbers so that very much changes the tenure of how people behave and how they react' (p. 320).

Teachers furthermore observed perverse incentives, such as to the desire to withdraw 'risky' students and retain 'marginal' students until the end of the academic year (to maximize funding units). All this effected their intrinsic motivation to educate. Even administrators felt the pressure: 'We're now in a position where we're so regulated – we get inspection, external verification, audit, all the bodies coming in and the pressure is felt throughout the organization both on the curriculum and noncurriculum sides' (p. 321).

Assignment

- 1. Consider your own position in your university. Do you feel that you have enough freedom to develop autonomously as a student? Have you perceived of any 'perverse incentives' in your educational program?
- 2. Do you agree with Kennedy that universities should strive to emphasize the social impact of education? What would engage you?
- 3. Start a fire.

Suggested Reading

For an orientation on the relationship between universities, politics, and economics, we recommend *Through a Darkly Glass: The Social Sciences Look at the Neoliberal University* (Magaret Thornton 2015). *Psychology in Crisis* by psychologist Brian Hughes (2018) grapples with some of the most fundamental problems in his field: theoretical sectarianism, psychology's susceptibility to produce irreproducible results, and convenient sampling, among others. The *Open Science Collaboration* project (2015) is a must read for anyone interested in the replication crisis.

References

- Barnard-Brak, L., Saxon, T. F., & Johnson, H. (2011). Publication productivity among doctoral graduates of educational psychology programs at research universities before and after the year 2000. *Educational Psychology Review*, 23(1), 65–73. https://doi.org/10.1007/ s10648-010-9146-3.
- Bornmann, L., Leydesdorff, L., & Van den Besselaar, P. (2010). A meta-evaluation of scientific research proposals: Different ways of comparing rejected to awarded applications. *Journal of Informetrics*, 4(3), 211–220. https://doi.org/10.1016/j.joi.2009.10.004.
- Bouter, L. M. (2015). Commentary: Perverse incentives or rotten apples? Accountability in Research, 22(3), 148–161. https://doi.org/10.1080/08989621.2014.950253.
- Brall, C., Maeckelberghe, E., Porz, R., Makhoul, J., & Schröder-Bäck, P. (2017). Research ethics 2.0: New perspectives on norms, values, and integrity in genomic research in times of even scarcer resources. *Public Health Genomics*, 20(1), 27–35. https://doi.org/10.1159/000462960.
- Canary, H. E. (2007). Teaching ethics in communication courses: An investigation of instructional methods, course foci, and student outcomes. *Communication Education*, 56(2), 193–208. https://doi.org/10.1080/03634520601113660.
- Ceci, S. J., Ginther, D. K., & Kahn, S. (2014). Women in academic science: A changing landscape. *Psychological Science in the Public Interest*, 15(3), 75–141. https://doi. org/10.1177/1529100614541236.
- Chambers, C. D. (2017). The seven deadly sins of psychology: A manifesto for reforming the culture of scientific practice. Princeton: Princeton University Press.
- Curnalia, R. M. L., & Mermer, D. (2018). Renewing our commitment to tenure, academic freedom, and shared governance to navigate challenges in higher education. *Review of Communication*, 18(2), 129–139. https://doi.org/10.1080/15358593.2018.1438645.
- Dal-Ré, R., Bouter, L. M., Cuijpers, P., Gluud, C., & Holm, S. (2020). Should research misconduct be criminalized? *Research Ethics*, 1–12. https://doi.org/10.1177/1747016119898400.
- Dobbie, D., & Robinson, I. (2008). Reorganizing higher education in the United States and Canada: The erosion of tenure and the unionization of contingent faculty. *Labor Studies Journal*, *33*(2), 117–140. https://doi.org/10.1177/0160449X07301241.
- Esposito, A. (2012). Research ethics in emerging forms of online learning: Issues arising from a hypothetical study of MOOCS. *The Electrotonic Journal of e-Learning*, 10(3), 315–325. http:// academic-conferences.org/ejournals.htm.
- Fanelli, D. (2010a). "Positive" results increase down the hierarchy of the sciences. *PLoS One*, 5(4), e10068. https://doi.org/10.1371/journal.pone.0010068.
- Fanelli, D. (2010b). Do pressures to publish increase scientists' bias? An empirical support from US states data. *PLoS One*, *5*(4), e10271. https://doi.org/10.1371/journal.pone.0010271.
- Fanelli, D., & Larivière, V. (2016). Researchers' individual publication rate has not increased in a century. *PLoS One*, 11(3), e0149504. https://doi.org/10.1371/journal.pone.0149504.
- Finkelstein, M. (2014). The balance between teaching and research in the work life of American academics. In J. Shin, A. Arimoto, W. Cummings, & U. Teichler (Eds.), *Teaching and research in contemporary higher education. The changing academy – The changing academic profession in international comparative perspective, vol 9* (pp. 229–318). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-6830-7_16.
- Garvin, D. (1980). The economics of university behavior. New York: Academic.
- Gelstein, S., Yeshurun, Y., Rozenkrantz, L., Shushan, S., Frumin, I., Roth, Y., & Sobel, N. (2011). Human tears contain a chemosignal. *Science*, 331(6014), 226–230. https://doi.org/10.1126/ science.1198331.
- Ginther, D. K., & Kahn, S. (2009). Does science promote women? Evidence from academia 1973-2001. In Science and engineering. Careers in the United States: An analysis of markets and employment (pp. 163–194). Chicago: The University of Chicago Press.
- Gračanin, A., Bylsma, L. M., & Vingerhoets, A. J. J. (2014). Is crying a self-soothing behavior? Frontiers in Psychology, 5, 502. https://doi.org/10.3389/fpsyg.2014.00502.

- Gračanin, A., van Assen, M. A. L. M., Omrčen, V., Koraj, I., & Vingerhoets, A. J. M. (2017a). Chemosignalling effects of human tears revisited: Does exposure to female tears decrease males' perception of female sexual attractiveness? *Cognition and Emotion*, 31(1), 139–150. https://doi.org/10.1080/02699931.2016.1151402.
- Gračanin, A., Vingerhoets, A., & Assen, M. A. (2017b). Response to comment on "Chemosignalling effects of human tears revisited: Does exposure to female tears decrease males' perception of female sexual attractiveness?". *Cognition & Emotion*, 31(1), 158–159. https://doi.org/10.1098/ rstb.2019-0262.
- Haggerty, K. D. (2004). Ethics creep: Governing social science research in the name of ethics. *Qualitative Sociology*, 27(4), 391–414. https://doi.org/10.1023/ B:QUAS.0000049239.15922.a3.
- Hangel, N., & Schmidt-Pfister, D. (2017). Why do you publish? On the tensions between generating scientific knowledge and publication pressure. *Aslib Journal of Information Management*, 69(5), 529–544. https://doi.org/10.1108/AJIM-01-2017-0019.
- Herbst, M. (2007). Alternative governance and management modes. In *Financing public universities*. *Higher Education Dynamics* (Vol. 18). Dordrecht: Springer. https://doi.org/10.1007/978-1-4020-5560-7_5.
- Hicks, D. (2012). Performance-based university research funding systems. *Research Policy*, 42(2), 251–261. https://doi.org/10.1016/j.respol.2011.09.007.
- Hood, C. (1991). A public management for all seasons? Public Administration, 69(1), 3–19. https://doi.org/10.1111/j.1467-9299.1991.tb00779.x.
- Hughes, B. M. (2018). Psychology in crisis. London: Macmillan.
- Huyer, S. (2015). Is the gender gap narrowing in science and engineering? In UNESCO science report: Towards 2030 (pp. 85–104). Paris: UNESCO publishing.
- Karkowski, A. M. (2010). Activities to help students appreciate committees that protect research participants. *Council on Undergraduate Research Quarterly*, 30(3), 11–15.
- Kezar, A., Maxey, D., & Holcombe, E. (2015). The professoriate reconsidered a study of new faculty model. The Delphi project on the changing faculty and student success, University of Southern California.
- Kuhn, T. S. (1970). The structure of scientific revolutions. Chicago: Chicago University Press.
- Löfström, E. (2016). Role-playing institutional academic integrity policy-making: Using researched perspectives to develop pedagogy. *International Journal for Educational Integrity*, 12(1), 5. https://doi.org/10.1007/s40979-016-0011-0.
- Makel, M. C. (2017). The empirical march: Making science better at self-correction. *Psychology of Aesthetics Creativity and the Arts*, 8(1), 2–7. https://psycnet.apa.org/doi/10.1037/14805-037.
- Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in psychology research. *Perspectives on Psychological Science*, 7(6), 537–542.
- Maturena, H., & Varela, F. (1980). Autopoiesis and cognition: The realization of the living (Vol. 42). Springer Science & Business Media.
- Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, 70(6), 487–498. https://doi.org/10.1037/a0039400.
- Merton, R. K. (1968). The Matthew effect in science: The reward and communication systems of science are considered. *Science*, 159(3810), 56–63. https://doi.org/10.1126/science.159.3810.56.
- Merton, R. K. (1988). The Matthew effect in science, II: Cumulative advantage and the symbolism of intellectual property. *Isis*, 79(4), 606–623. https://doi.org/10.1086/354848.
- Moustafa, K. (2014). The disaster of the impact factor. *Science and Engineering Ethics*, 21(1), 139–142. https://doi.org/10.1007/s11948-014-9517-0.
- Naimi, L. (2007). Strategies for teaching research ethics in business, management and organisational studies. *The Electronic Journal of Business Research Methods*, 5(1), 29–36.
- Ogbogu, C. O. (2011). Gender inequality in academia: Evidences from Nigeria. Contemporary Issues in Education Research, 4(9), 1–8. Retrieved from: https://files.eric.ed.gov/fulltext/ EJ1072918.pdf.
- Open Science Collaboration. (2015). Science, https://doi.org/10.1126/science.aac4716
- Parson, T. (1951). The social system. Glencoe: The Free Press.

- Resnik, D. B. (1998). The ethics of science: An introduction. London/New York: Routledge.
- Schwartz, J. M. (2014). Resisting the exploitation of contingent faculty labor in the neoliberal university: The challenge of building solidarity between tenured and non-tenured faculty. *New Political Science*, 36(4), 504–522. https://doi.org/10.1080/07393148.2014.954803.
- Siegel, D., & Baveye, P. (2010). Battling the paper glut. *Science*, *329*(5998), 1466. https://doi. org/10.1126/science.329.5998.1466-a.
- Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11), 1359–1366. https://doi.org/10.1177/956797611417632.
- Smith, N. C. (1970). Replication studies: A neglected aspect of psychological research. American Psychologist, 25(10), 970–975. https://psycnet.apa.org/doi/10.1037/h0029774.
- Sobel, N. (2017). Revisiting the revisit: Added evidence for a social chemosignal in human emotional tears. *Cognition & Emotion*, 31(1), 151–157. https://doi.org/10.1080/02699931.201 6.1177488.
- Sponholz, G. (2000). Teaching scientific integrity and research ethics. Forensic Science International, 113(1-3), 511–514.
- Stern, J. E., & Elliott, D. (1997). The ethics of scientific research: A guidebook for course development. Hanover/London: University Press of New England.
- Stone, D. (2002). *Policy paradox: The art of political decision making*. New York: W.W. Norton & Company.
- Sweet, S. (1999). Using a mock institutional review board to teach ethics in sociological research. *Teaching Sociology*, 27(1), 55–59. https://www.jstor.org/stable/1319246.
- Thornton, M. (2015). *Through a darkly glass: The social sciences look at the neoliberal university*. Acton: ANU Press.
- Van Bavel, J. J., Mende-Siedlecki, P., Brady, W. J., & Reinero, D. A. (2016). Context sensitivity in scientific reproducibility. *Proceedings of the National Academy of Sciences*, 113(23), 6454–6459. https://doi.org/10.1073/pnas.1521897113.
- van den Besselaar, P., & Leydesdorff, L. (2009). Past performance, peer review and project selection: A case study in the social and behavioral sciences. *Research Evaluation*, 18(4), 273–288. https://doi.org/10.3152/095820209X475360.
- Van Wesel, M. (2016). Evaluation by citation: Trends in publication behavior, evaluation criteria, and the strive for high impact publications. *Science and Engineering Ethics*, 22(1), 199–225. https://doi.org/10.1007/s11948-015-9638-0.
- von Unger, H. (2016). Reflexivity beyond regulations. Teaching research ethics and qualitative methods in Germany. *Qualitative Inquiry*, 22(2), 87–98.
- Wiener, H. (1965). *Cybernetics, second edition: Or the control and communication in the animal and the machine.* Cambridge: The MIT Press.

References for Case Study: How to Start a Fire

- Biesta, G. (2007). Why "what works" won't work: evidence-based practice and the democratic deficit in educational research. *Educational Theory*, 57(1), 1–22. https://doi. org/10.1111/j.1741-5446.2006.00241.x.
- Boocock, A. (2013). Further education performance indicators: a motivational or a performative tool? *Research in Post-Compulsory Education*, 18(3), 309–325. (2013. https://doi.org/10.108 0/13596748.2013.819272.
- Kennedy, J. (2016). An educated guess. In M. Flikkema (Ed.), Sense of serving. Reconsidering the role of universities today (pp. 32–41). Amsterdam: University Press.
- Marczyk, J. (2017, March). Academic perversion. *Psychology Today*. Retrieved from https://www. psychologytoday.com/us/blog/pop-psych/201703/academic-perversion

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), 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 license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license 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.

