From the Researcher to the Integrity of Knowledge Production

Sven Ulpts

From the Researcher to the Integrity of Knowledge Production

PhD Dissertation

Politica

© Forlaget Politica and the author 2026

ISBN: 978-87-7335-357-8

Cover: Svend Siune
Print: Fællestrykkeriet, Aarhus University
Layout: Annette Bruun Andersen

Submitted May 6, 2026
The public defence takes place August 28, 2026
Published August 2026

Forlaget Politica
c/o Department of Political Science
Aarhus BSS, Aarhus University
Bartholins Allé 7
DK-8000 Aarhus C
Denmark

Table of Contents

Acknowledgements 9

Chapter 1: Introduction 11
Tensions between policy and research 11
The need for an emic perspective 14
This dissertation: object, setting, and scope 15
Research aims and research questions 15
Research integrity, ethics, preregistration, replication, and reproducibility 16
The contribution 17
Structure of the dissertation 18

Chapter 2: Situating research integrity 21
A short history of research integrity 21
Research integrity and the emergence of Open Science 25
The concept research integrity 28
Researchers' perspective on research integrity 33

Chapter 3: Theoretical framework: the transepistemic arena of 'good research' 39
Credibility cycle 43
Credibility work 45
Application of the framework in the dissertation 47

Chapter 4: The ethnographic methodology 49
Motivation 49
Ethnographies of science 51
Meetings as an entry point 55
Epistemology 56
Positionality statement 58
Data collection 58
Getting access and entering the field 60
Field observations 61
Interviews 62
Analysis 65
Ethical considerations 65

Chapter 5: Research integrity from within: good science in the lab 69
The moral scholar and personal integrity 73
Taking responsibility and owning the research 76
Honesty 79
Humility and critical reflexivity 81
The role of the research system 85
The researcher in a group of researchers 88
The integrity of the research 93
Misconduct: fabrication, falsification, and plagiarism. 93

Methodological rigour 95
Hypothesis testing 96
Research ethics ('do no harm') 96
Prevention of bias 97
Societal relevance and practical utility 98
The relationship between research integrity and Open Science 99
Discussion and conclusion 106
Chapter 6: Ethics in the lab 111
Introduction 111
The ethics of dealing with participants 114
The ethical colleague 120
Discussion and conclusion 127
Chapter 7: Preregistration in the lab 129
Introduction 130
Critical perspectives on the practice and valuation of preregistration 132
Researching preregistration 134
Preregistration talk in the lab 137
Thinking through your study with preregistration 137
Preregistration to constrain researcher degrees of freedom 138
A timestamped firewall between confirmation and exploration 141
Preregistration as a credibility management strategy 146
Practicalities and problems with practicing preregistration 149
The uncertainty of piloting 149
Time investment and workload for preregistration 153
Preregistration is maybe not a prison but a punishment 156
Overwhelming students with preregistration 163
Preregistration: a story of credibility and hardship 165
Chapter 8: The practices of replication and reproducibility in the lab 169
Introduction 169
Science studies on the practices of reproducibility and replication 174
Replication and reproducibility talk in the lab 177
Meaning(s) of replication and reproducibility in the lab 178
Confusing terms 182
Function of reproducibility and replication 183
Providing a stable foundation for incorporation 184
Stories of their own practice 187
Failed replications 190
Reproducibility and replication and a perceived demand for novelty 193
Reproducibility and replication in the lab 197

Chapter 9: Discussion and conclusion 201
Good research from within: the emic perspective on research integrity 203
With lab studies from science studies over credibility to research integrity 208
Limitations 211
Conclusion 212
Summary 215
Dansk resumé 217
References 219
Appendices 261
Appendix A 261
Appendix B 265

Acknowledgements

In contrast to the rest of my dissertation, I'll keep this short. Let's first get the usual suspects out of the way. I, of course, thank my family – Jutta, Rudolf, Lars, Nils and Hermine – for their unwavering support and open ears when I complain about life and academia. With my family out of the way, it's time for the most important acknowledgment before the usual academic formalities: thank you, Njall, for all the free food!

I would like to thank my first supervisor Lise for providing me the entry point to academia by hiring me in the first place. Although I am unsure she knew what she was getting herself into by hiring me and taking on the burden of being my supervisor – an unsupervisable PhD student. Special thanks to Jesper! When Lise left CFA for a while, Jesper, knowing what he was getting himself into, took over from Lise and switched roles from second to first supervisor, full of 'enthusiasm'. Furthermore, I much appreciate the assessment committee – Tine, Martin, and Willem – for reading my dissertation twice and providing helpful feedback, criticism and suggestions for improvements. Now, to the person who's been my office mate for the longest time at CFA, thanks (young) Mads. I would also like to thank Alan for the fun conversations about the old times of science studies (and some small corners of STS).

For my first research stay abroad, I visited Jochen Gläser's group at the Department for Social Studies of Science and Technology at the TU in Berlin, Germany. Thanks, Jochen, Grit, Markus, Chris, Tobias, Suse, and Clara, for providing me an introduction to the sociology of science and the fun times, especially our obligatory escape room expeditions. Roughly a year later, I went back to Berlin to visit the Robert K. Merton Center for Science Studies. Thank you to the team at the RMZ for hosting me and letting me be a part of their vibrant research culture as well as engaging in thought-provoking discussions. I won't mention everybody, but special mention to Martin, Stephan and Sheena. Moreover, I would like to thank Sheena for the ongoing collaboration and the much-needed distractions from my PhD.

Lastly, I am immensely grateful to the researchers that let me into their workplaces and introduced me to their work routines and made me feel welcomed in their labs as well as took the time to answer my questions. Without them none of this research would have been possible in the first place.

Chapter 1: Introduction

Over the last two decades the impression of a 'crisis in scientific integrity' seems to have emerged (Amorim & Bacchi, 2025, p. 1), with research integrity becoming an increasingly visible topic in science policy, institutional governance, and scholarly debate. Concerns about scientific credibility, questionable research practices, research misconduct, reproducibility, transparency, and resulting calls for reforms have contributed to a situation in which research integrity is no longer a marginal issue or only limited to specific disciplinary contexts, but an influential category through which science is investigated, discussed, evaluated, and governed (Aubert Bonn & Pinxten, 2019; Carcausto-Calla et al., 2023; Field et al., 2024; Hesselmann & Reinhart, 2024; NASEM, 2017; Tang, 2025b; Vendé et al., 2025; Zuckerman, 2020). This is visible not only in scholarly debate, but also in codes of conduct, responsible conduct training initiatives, research integrity officers, and broader reform agendas, like those associated with Open Science (OS) (Abbott, 2019; ALLEA, 2023; Bouter, 2024; DFG, 2019; Horbach et al., 2024; KNAW, 2018; UKRIO, 2025; Winchester, 2018). While these developments have mobilised substantial investments and reforms aimed at ensuring as well as improving research integrity, studies also reveal deep conceptual ambiguity and context dependency about what research integrity actually means and how it is enacted in practice across research communities (see e.g., Davies & Lindvig, 2021; Horbach & Halffman, 2017; Kerkhoff & Overeem, 2021; Penders et al., 2009; Pols et al., 2024; Roberts et al., 2020; Shaw, 2019). At the same time, previous studies suggest that there may be a disconnect between these official and scholarly understandings of research integrity and the ways in which researchers themselves understand and enact 'good research' in practice. While research integrity is increasingly defined, promoted, and regulated, less is known about how researchers at local sites of knowledge production perceive these ideas, how they negotiate them in their own work, and what actually motivates the practices associated with 'good research'. This is what this dissertation aims to address, based on a multisite lab ethnography of six psychology, cognitive science and neuroimaging laboratories in Germany, the UK, and the Netherlands.

Tensions between policy and research

Although research integrity has become an important scholarly and 'policy object' (Davies & Lindvig, 2021; Sin, 2014), there are critical indications that

official and scholarly understandings do not map neatly onto researchers' own values and everyday practices. Previous studies attempting to capture researchers' perceptions, experiences, and understandings of research integrity indicate that there are areas of disconnect and misalignment between the researchers' values, experiences, and perceptions, on the one side, and science policy as well as the community dedicated to investigating research integrity, on the other. (see e.g., Brenninkmeijer & Pols, 2026; Davies, 2019; Davies & Lindvig, 2021; De Vries et al., 2006; Hangel & Schickore, 2017; Horbach & Halffman, 2017; Pols et al., 2024; Schickore & Hangel, 2019; Valkenburg et al., 2021).

Claims of misalignments between science regulation and researchers' values and conduct, are not new. Two recent and related examples of discourse about such misalignments are Open Science and metrics-based research assessment. The first is a consequence of calls for and implementation of recent science reforms, while the second is a reason behind calls for reform. There is an ongoing discourse about perceived misalignments between research evaluation by science policy and governance based on quantitative metrics and what researchers themselves see as high-quality research according to their community's internal preferences and lived experiences (see e.g., Benedictus et al., 2016; Biagioli & Lippman, 2020; Gläser, 2024; Robinson-Garcia et al., 2023; Rushforth & Hammarfelt, 2023). In the context of Open Science, which in recent years has become part of the research integrity conversation with topics such as replication, reproducibility, data sharing, transparency, and preregistration (see e.g., Bouter, 2024; Haven et al., 2022; Tang, 2025b), a discourse is emerging about potentially detrimental disconnects and misalignments with diverse and local epistemic communities, because incentivised or required Open Science practices and standards are seen as unfeasible or irrelevant for some epistemic contexts (see e.g., Bazzoli, 2022; Cole et al., 2024; Devezer & Penders, 2023; Guttinger, 2020; Horbach et al., 2026; Hostler, 2023; Leonelli, 2018, 2022, 2023, Lilja, 2021, 2025; Ma & Davin, 2025; Manco, 2024; Moors, 2026; Patton, 2024; Penders et al., 2019, 2020; Ulpts & Schneider, 2025). In the case of research integrity specifically, some conceptual work has also already been done to frame apparent misalignments (Labib et al., 2024; Valkenburg et al., 2021; Vincent & Hooper, 2025).

Valkenburg and colleagues (2021) distinguish between aspirational values of researchers that are embedded in their community specific (epistemic) culture and practices on the one side, and sanctionable values related to research integrity which are part of institutional research governance, courses, and guidelines, on the other (Knorr-Cetina, 1999; Valkenburg et al., 2021; also see Horbach and Halffman, 2017). According to Valkenburg and co-authors (2021), sanctionable values, of the bureaucracies and protocols, are trans-

parency, fair credit, human dignity, and the prevention of plagiarism, falsification and fabrication. They list integrity, inquisitiveness, reflexivity, collegiality and trust as important aspirational values that are anchored in the actual practices and discourse of researchers within their epistemic cultures (Valkenburg et al., 2021).

Based on their distinction between commitment-based and rule-based forms of research integrity, according to Labib and colleagues (2024), the proper functioning of research governance towards research integrity requires that research policy and governance align with the understandings, practices, and values within the research communities. Otherwise, research regulation and governance may 'impinge on', or 'colonize' research communities by 'replacing' their internal values and practices (Labib et al., 2024, p. 925).

Vincent and Hooper (2025) speak of a deontic shift and a process of thinning of values during the codification of 'good research' in codes of conduct, as wider applicability and communication to larger and diverse audiences across research communities and stakeholders, simplification and translation become a necessity. They identify three problems with thin values: 1) mistaking the map for the territory, 2) proxy failure, and 3) crowding out and nudging. Similar to ethical approvals not (necessarily) making research ethical, codes of conduct as maps can only provide orientation, but actual good research practices emerge and are enacted within specific contexts of the researchers' territories (Vincent & Hooper, 2025). Due to the need for monitoring, assessment and enforcement in the science system, research integrity governance and evaluation often relies on metrics and indicators as proxies and, therefore, run the danger to be corrupted and misused (Campbell, 1979; Vincent & Hooper, 2025). Although well-intentioned, codes of conduct, incentives and requirements for good research practices might 'distract from what makes good research valuable' by detaching them from the thick ethos of the specific contexts and epistemic cultures they are enacted in (Vincent & Hooper, 2025, p. 11). As, for instance, ethical considerations of openness and privacy in qualitative research illustrate (see e.g., Moors, 2026), values of proper conduct can conflict with each other (also see Helgesson & Bülow, 2023; Peels et al., 2019; van Drimmelen et al., 2025). Crowding out can occur as the process where external incentives and rewards replace or push away intrinsic motivation to demonstrate research integrity by conducting 'good research' or acting ethically (also see cognitive dissonance, Festinger, 1957).

Therefore, the issue is not only conceptual ambiguity (Kolstoe & Pugh, 2024; Shaw, 2019; Steneck, 2006), but also that governance, reform, and guidance may be built on assumptions about 'good research' that are insufficiently grounded in how researchers themselves understand and enact it in practice. Hence, there are likely cases where research conduct does not

correspond with notions of 'good research' that can be found in the research integrity literature and research policy, not because there is a lack of education, those are bad researchers, or incentives pressured them to misbehave, but because their community's internal preferences, values, and research conditions do not align with them. To contribute to that literature and critical discussion, I set out to, further investigate researchers' perceptions, understandings, experiences, and negotiations of research integrity in addition to the motivations that underlie their enactment of practices that have been related to research integrity from within laboratories (see e.g., Brenninkmeijer & Pols, 2026; De Vries et al., 2006; Gläser, 2024; Hangel & Schickore, 2017; Laudel, 2024; Müller et al., 2024; Pols et al., 2024; van Drimmelen et al., 2024; 2025).

The need for an emic perspective

While some argue that current endeavours to improve research integrity and associated reforms are working (see e.g., Korbmacher et al., 2023; Nelson et al., 2018; Protzko et al., 2023; Vazire, 2018), others question their working, suggesting a need to look beyond just perverse incentives or bad scholars and instead look at the various realities and conditions of everyday research from within diverse research communities (Nelson, 2021; Peterson, 2021)[1]. Furthermore, although normative guidance in the form of protocols, teaching and regulation concerning responsible conduct, 'good science', or research integrity do exist, researchers still have extensive discretion or researcher degrees of freedom surrounded by uncertainty, when making decisions about steps and practices in their various research processes (Gelman & Loken, 2014; Knorr-Cetina, 1983; Simmons et al., 2011; van Drimmelen et al., 2024, 2025). Moreover, the literature suggests that, to a certain extent, researchers will inevitably navigate around regulatory structures and restrictions to do their work (see e.g., De Vries et al., 2006; Elizondo & Kaltenbrunner, 2024; van Drimmelen et al., 2025), suggesting that degrees of freedom and discretion are not eliminated by protocols and regulation, but just moved elsewhere, while many decisions remain at the researchers' discretion, tacit, or invisible, even to the researchers themselves (see e.g., van Drimmelen et al., 2024).

If research integrity is to be understood as more than a policy label, list of formal norms, or topics to look out for (Labib et al., 2021; Sørensen et al., 2021), then it must be studied from within the settings in which research is actually conducted, discussed, supervised, and evaluated (see e.g., Knorr-Cetina, 1995; Nelson, 2021; Peterson, 2021; Pols et al., 2024).

[1] The Protzko and colleagues (2023) article has been retracted in the meantime (Protzko et al., 2024).

Methodologies like interviews, surveys, and focus groups alone mostly provide an etic perspective, because those require translations to outsiders, who investigate 'good' or 'bad research' (Gläser, 2024). Ethnographies of scientific practice allow to capture an emic perspective on research integrity from the inside, as the enactment of 'good research' according to the researchers' preferences and the specific conditions of the local sites of knowledge production (Gläser, 2024; Knorr-Cetina, 1995; Pols et al., 2024; Stephens & Lewis, 2017). As such an ethnographic methodology with a combination of field observations and follow-up interviews provides us an inside look into how researchers themselves understand and negotiate 'good research' in practice within their laboratories as well as the context of the wider research system.

This dissertation: object, setting, and scope

I aim to provide an emic perspective on research integrity, by capturing researchers' notions of research integrity as forms of 'good research' from within their sites of knowledge production. To that end, inspired by classical lab ethnographies from Lynch (1985), Latour and Woolgar (1979), Knorr-Cetina (1981, 1999) and Traweek (1988), I present an ethnographic study of the experiences, understandings and perceptions of the concept of research integrity from researchers from within their psychology, cognitive science, and neuroimaging laboratories in Germany, the United Kingdom and the Netherlands. The ethnographic study is based on field observations during lab meetings, supervision meetings and journal clubs to capture how scholars talk about and negotiate forms of 'good research' during such situations of evaluation and negotiation (Laudel, 2024; Müller, 2024; Philipps & Paruschke, 2024). The field observations are followed up on and complemented by semi-structured interviews informed by the themes that emerged during the interactions between scholars in the meetings. As the focus is more on idiographic insights from researchers inside specific sites of knowledge production within the selected laboratories, I do less aim to make general claims across all the diverse research contexts in all the sciences or geographic locations (Kreimer, 2022).

Research aims and research questions

I aim to complement the existing research integrity literature concerning the conceptual ambiguity and potential misalignments between research communities' internal practices, conditions, and preferences concerning 'good research', on the one side, and research policy and research on research integrity, on the other. To that end I present an emic perspective of research integrity from within research communities by addressing the following primary

research question: How do researchers understand and enact research integrity or 'good research' in their everyday activities. This overarching question will be answered by focusing on the two secondary questions:

1. How do researchers talk with each other about and negotiate values and practices that have been linked to research integrity in the laboratory settings?
2. What motivations, concerns, and practical conditions shape these understandings and practices?

The overarching research question will be answered by the ethnographic combination of field observations and interviews. The field observations during meetings lay the foundation for addressing the research questions while situating the researchers' discourse and reflections about 'good research' within the context of their laboratories. The field observations mostly address the first sub-question by capturing the themes that determine how researchers negotiate 'good research' as part of their lab internal discourse. The interviews, which build on the themes that emerged during the field observations, open the dialogue with the researchers to reflect on their understandings, perceptions, and experiences of 'good research' as well as their motivations and practical conditions behind the practices they associate with 'good research'. As such, the interviews aid in answering the first sub-question, but especially address the second sub-question regarding motivations, concerns and practical conditions that shape their understandings and practices.

Research integrity, ethics, preregistration, replication, and reproducibility

Research integrity and research ethics are inherently linked, not just due to their history as regulatory concepts for proper research conduct in reaction to cases of research practices that were labelled as improper, but also institutionally, conceptually and practically for scientific knowledge production, especially in research involving human or animal subjects (Kolstoe & Pugh, 2024; Mickel et al., 2023; Moors, 2026). Moreover, in the context of current crisis declarations and resulting reform movements, the heavy circulation of Open Science concepts within the research integrity discourse, and the often assumed mutually reinforcing relationship between Open Science and research integrity (see e.g., Bouter, 2024; Haven et al., 2022; Tang, 2025b), issues surrounding replication, reproducibility, and preregistration occupy a central place in current research integrity discourse. Reproducibility and replication have a long history of being linked to research quality and integrity (Francis, 1989; NAS et al., 1992; Steneck, 2007) and in the recent literature they are

presented as critical concerns about research integrity (Diaba-Nuhoho & Amponsah-Offeh, 2021; Resnik & Shamoo, 2017). The reporting practice of preregistration is one of the proposed interventions in the social and cognitive sciences to ensure and improve research integrity by inducing more transparency and accountability into the practices and communication of research (Emons et al., 2025; Flis, 2022; Penders, 2022). Furthermore, ethical considerations, preregistration, replication and reproducibility regularly came up as critical issues for 'good research' in interactions between researchers during the field observations. As such they help us to go beyond abstract notions of 'good research' and research integrity by exploring the actual enactments of practices linked to 'good research' (Valkenburg et al., 2021). Thus, they further provide a window into how different notions of 'good research' are grounded in the researchers' contexts and lived experiences as well as how they relate to each other. Therefore, following the chapter on the researchers' meanings, understandings and perceptions of research integrity I will explore how the enactments of research ethics, preregistration, reproducibility, and replication relate to their understandings, experiences and perceptions of 'good research'.

The contribution

The dissertation provides an ethnographic account of how researchers themselves understand, negotiate and enact 'good research' within research communities in everyday laboratory settings. Therefore, it provides further material to address the conceptual ambiguity of research integrity, and it can help to clarify where alignments and misalignments may arise between policy discourse, reform agendas, research on research integrity, research practices, and research conditions. As such it can further contribute to ongoing discussions about how diverse forms of research integrity depend on the epistemic properties and contexts of the research at hand (Penders et al., 2009). Due to the selection of psychology related disciplines for fieldwork, which together with biomedicine build the epicentre of metascience and current Open Science reforms (Nelson et al., 2021; Peterson & Panofsky, 2023), reflections on Open Science themes like reproducibility, transparency and preregistration further allow for exploration of the relation between Open Science and research integrity. Additionally, the inclusion of two chapters on practices provides a perspective to investigate the enactments and negotiations of forms of 'good research' in practice, by going beyond the abstract notions of research integrity and research ethics (Pickering, 1992). By building theoretically and methodologically on established science studies work on credibility, and laboratory studies specifically, the dissertation also gives an opening for dialogue between science studies and research on research integrity. All in all the dissertation contributes conceptually by bringing this material into dialogue with

research integrity scholarship and science studies as well as exploring the relationship between Open Science, research ethics, and research integrity, and practically by clarifying where alignments and misalignments may arise between policy discourse, reform agendas, research on research integrity, on the one side, and research practice in their local sites of knowledge production with certain conditions and internal preferences for 'good research', on the other.

Structure of the dissertation

In the next chapter I will situate research integrity in the larger context by first providing a short outline of its historical and political origin, then I will touch on how the current movement towards Open Science relates to research integrity, followed by an overview of recent work on clarifying the concept of research integrity, culminating in a short review of existing research into researchers' experiences and perceptions of research integrity. Afterwards, I will present my theoretical framework in Chapter 3. I situate research integrity in the larger transepistemic arena and within the credibility cycle (Knorr-Cetina, 1982; Latour & Woolgar, 1979) as a component of research policy entering research communities mainly through research evaluation. In Chapter 4 I will outline my ethnographic methodology and provide a short background about what lab ethnographies are as well as motivate why this methodology provides an emic approach to research integrity as an entry point to the researchers' perceptions and experiences concerning research integrity as well as related practices. Chapter 5 will be the first empirical chapter. In it I present the researchers' conceptualisations, experiences and perceptions of research integrity by mainly addressing the first sub-research question of how researchers talk with each other about and negotiate values and practices related to research integrity. The aim is to capture what research integrity as 'good research' means to the researchers in the context of their day-to-day research situations. In the following Chapter 6, I will inquire into how researchers relate the ethical considerations of their conduct to what constitutes 'good research'. Hence, this chapter, to a certain degree, addresses the secondary research questions one and two. The last two empirical chapters, seven and eight, will focus on the experiences and perceptions of preregistration, replication, and reproducibility practices. In Chapter 7, I explore the researchers' experiences with using the reporting practice of preregistration, more seldomly registered reports, as a tool to acquire, convey, and maintain credibility in their respective epistemic communities (Daston, 1995; Meskus et al., 2018; Latour & Woolgar, 1979; Li Vigni et al., 2023; Shapin, 1984, 1995a; Shapin & Schaffer, 1985). In the last empirical Chapter 8, I will first outline what the researchers actually understand replication and reproducibility to mean.

Then, in the main part of this chapter, aided by sociology of scientific knowledge literature on replication (see e.g., Ashmore, 1988; Brenninkmeijer et al., 2025; Collins, 1975, 1985, 2016; Mulkay, 1984; Mulkay & Gilbert, 1986; Shapin & Schaffer, 1985; Travis, 1981), I will explore their practices and motivations behind them to understand what utility they have for the researchers by fulfilling specific epistemic functions in their management of credibility (also see Brenninkmeijer et al., 2025; Guttinger, 2018, 2019; Peterson & Panofsky, 2021). To round off this dissertation, in Chapter 9 I will provide a general reflection of my ethnographic work and experiences by summarising and discussing the main findings and situating them within the research integrity literature. To that end, I will further explore how they relate to previous research, indicating tensions between research realities and research integrity regulation, as well as discuss how long-established science studies fit into the current research integrity discourse. Furthermore, I will elaborate on what the presented findings might mean for research integrity policy and the research on research integrity community.

Chapter 2: Situating research integrity

In this chapter, I aim to situate the research integrity discourse within the larger context of regulating and reforming science. I will first provide a short historical overview of how research integrity emerged as a regulatory concept in reaction to public outcry over cases of research misconduct in the biomedical sciences and travelled from the US to Europe. Afterwards, I will shortly introduce the emergence of Open Science in the research integrity literature, to situate it in the context of other research communities dedicated to understanding and reforming science. Then I will give a small review of contests and clarifications concerning the meanings and conceptualisations of research integrity, as well as how it relates to other concepts like research quality and research ethics. I will end this chapter with an outline of recent research into researchers' own perceptions and experiences of research integrity from within their research communities.

A short history of research integrity

Preceding the regulation of proper conduct of science in the name of research integrity, was the ethical regulation of research conduct with human participants in reaction to public cases of gross mistreatment of humans. Indicating that the regulation of proper research conduct is naturally a reactionary phenomenon, as the history of ethical regulation in medical research demonstrates as well. For example, as part of the verdict in the Nuremberg trials, the inhumane medical experimentation as part of the Nazi war crimes lead to the Nuremberg code of 1947, containing ten features of acceptable research, followed by the Helsinki Declaration of 1964 as the first official guiding document of its kind regulating biomedical research (Dhai, 2014), and lastly the National Research Act of 1974 of the US Congress in reaction to public outcry and show of distrust in science following the revelations of the Tuskegee Syphilis study (1932-1972), culminating in the Belmont report of 1979 one of the foundations for the widespread dependence on institutional review boards (Moors, 2026; Nagai et al., 2022).

Similarly, research integrity entered the stage as a policy and regulatory concern following high-profile misconduct cases in the 1970s and 1980s US context. The US Office of Scientific Integrity (OSI) and the Office of Scientific Integrity Review (OSIR) were founded in 1989, reacting to widely publicised and scandalised cases of research misconduct in the US, like the Baltimore affair (Lock, 1999), that were the topic of heated discussions in US Congress

and the media alike (Guston, 2000; Hesselmann & Reinhart, 2024; Pascal, 1999; Price, 2013). In 1992 the Office of Research Integrity (ORI) replaced OSI and OSIR to take over the responsibility of dealing with cases of misconduct in research funded by the US National Institute of Health (NIH) to make the process more independent as well as formal, to allow for due process of the accused parties and take the workload off the funders (see e.g., Price, 2013; Guston, 2000; Hesselmann & Reinhart, 2024), but also, apparently, to direct the attention from congress and the public away from the NIH (Price, 2013). ORI became responsible for tasks like, defining misconduct, the handling of misconduct allegations, oversight of policy compliance by research institutes, like responsible conduct of research teaching mandates, and the protection of whistleblowers. Hence, the introduction of research integrity and the associated regulatory body into science was the result of an intervention of political or governmental oversight. In other words, unlike many current Open Science and research integrity reform endeavours, originally the notion of research integrity and the establishment of this regulatory body originated from outside the strictly scientific community, with political actors often perceived as infringing on academic freedom (Hesselmann & Reinhart, 2024; Kim & Park, 2013). Interestingly, research integrity as an official issue in science and its institutionalisation have emerged in the biomedical fields (Guston, 2000; Pascal, 1999; Price, 2013; Roberts et al., 2020), a heritage that still affects research integrity policy and scholarship alike (see Clancy et al., 2026 for problematic nature of biomedical assumptions inherent to research integrity framings; see Peršak, 2025 and UK Committee on Research Integrity, 2024, p. 5 for 'quantitative bias' in research integrity). This parallels the disciplinary origin of metascience and meta-research in the 2000s and 2010s (see e.g., Ioannidis, 2005; Nelson et al., 2021, 2022). Importantly, at the time, the scientific community largely rejected and argued against the intrusion of politics, policy and more generally other non- or extra-scientific regulatory bodies' attempts to control or steer their conduct in the name of 'good', 'better' or 'legitimate science', because they remained convinced that (natural) science is self-correcting and that the portrayal of research misconduct in the media was exaggerating the prevalence of such issues (see e.g., Felt & Frantz, 2026; Hesselmann & Reinhart, 2024; Price, 2013; Steneck, 1994; Tamot et al., 2013). In Merton's words, research misconduct was seen as 'extremely infrequent' (Merton, 1957, p. 651). According to Bimber and Guston (1995), the developments set in motion by the US congress leading to the emergence of research integrity as a regulatory concept with the founding of ORI can be seen as 'the deconstruction of scientific exceptionalism', and it practically 'institutionalized the recognition that the Mertonian normative structure [of science] is incomplete and that the market of ideas needs regulation' (pp. 565-567). This

regulatory intervention did not just usher in a process of formalisation and standardisation concerning the handling of misconduct cases, but also visible, still ongoing, boundary work surrounding scientific autonomy and the negotiation of the meaning of legitimate scientific practice as well as what the scientific community actually is (Bezuidenhout & Beeler, 2021; Felt & Frantz, 2026; Gieryn, 1983, 1995, 1999; Guston, 2000; Knorr-Cetina, 1982; Sismondo, 2025; Steneck, 1994).

This biomedical origin and sense of attack on scientific self-regulation was also, more or less, mirrored in the European context. Following the examples set in the US to varying degrees. With Scandinavia leading the way, Europe also started to form numerous research integrity regulatory bodies and implement related policies (Bosch, 2010; Desmond & Dierickx, 2021; Videnoja et al., 2024). For example, in Germany, after the widely scandalised misconduct case of the cancer researcher Firedhelm Herrmann in 1997 (Hagmann, 2000), the German Research Foundation (DFG) published a memorandum in 1998 entitled 'Safeguarding good scientific practice' (*'Sicherung guter wissenschaftlicher Praxis'*) (DFG, 1998/2026). As Abbott (1999) put it, with the Friedhelm Herrmann case 'German science lost its innocence' (p. 765), shedding doubt on science's ability to regulate and correct itself (Tuffs, 2000). In response to the 1998 memorandum and its recommendation 16 '*Ombudsman for Science*' in 1999 the German Ombudsman system was founded (von Bargen, 2013). While being inspired by the US Office for Research Integrity, the memorandum explicitly states that 'it should not have a mandate to conduct its own investigations like, for instance, the Office of Research Integrity of the US Public Health Service' (DFG, 1998/2026, p. 67). The implementation of the German Ombudsman system is necessarily different from the US Research Integrity politics due to differences in the relationship between politics and science as regulated by national laws. Germany's article five paragraph three of the basic law is supposed to protect academic freedom and limit political interference into the sciences, for instance, by securing the autonomy of higher education institutes when it comes to handling suspected cases of misconduct (von Bargen, 2013; Videnoja et al., 2024). The memorandum was updated in 2013 with 17 instead of 16 recommendations and in 2018 the DFG formulated the first code of conduct for research integrity in Germany which was implemented in the summer of 2019 (Hochbruck, 2024).

Since historically research integrity, as a regulatory concept, emerged in the context of the Office of Research Integrity (ORI), which was largely founded to deal with research misconduct cases, and research integrity is still often conceptualised as the opposite of improper research conduct (Horbach & Halffman, 2017; Horbach et al., 2024), it might be insightful, for understanding what research integrity actually entails in this rather negative sense,

to see what the definition of research misconduct was at the time and how it changed during the setup of that office: 'Misconduct or Misconduct in Science means fabrication, falsification, plagiarism, or other practices that seriously deviate from those that are commonly accepted within the scientific community for proposing, conducting, or reporting research. It does not include honest error or honest differences in interpretations or judgments of data' (US Public Health Service, 1989, p. 32447). Then, the 1992 report on *Responsible Science: Ensuring the Integrity of the Research Process*, also provided a definition for research integrity, which was distinguished from the integrity of science and presented as 'the adherence by scientists and their institutions to honest and verifiable methods in proposing, performing, evaluating, and reporting research activities' (NAS et al., 1992, pp. 24-25). The report also identified three aspects of concern that need to be addressed to ensure the integrity of the research process(es): 1) misconduct in science, 2) questionable research practices and 3) other misconduct. The definition of misconduct in this report does not include other serious deviations, because it was seen as too vague for a federal definition (NASEM, 2017). In the report, 'Misconduct in science is defined as fabrication, falsification, or plagiarism, in proposing, performing, or reporting research. Misconduct in science does not include errors of judgment; errors in the recording, selection, or analysis of data; differences in opinions involving the interpretation of data; or misconduct unrelated to the research process' (NAS et al., 1992, p. 27). Questionable research practices are defined as 'actions that violate traditional values of the research enterprise and that may be detrimental to the research process. However, there is at present neither broad agreement as to the seriousness of these actions nor any consensus on standards for behavior in such matters' (NAS et al., 1992, p. 28). As indicated above, the lack of agreement has not changed much in the following years (see e.g., Ravn & Sørensen, 2021). Interestingly, it is stated that questionable research practices are not considered misconduct because they do not directly damage the integrity of the research, but they need to be acknowledged because they, for instance, affect the confidence in research, waste time, deviate from research traditions' practices, and can support an environment that fosters misconduct. Other forms of misconduct are forms of misconduct that are not unique to science, like sexual harassment and the misuse of funds. Potential causes for research misconduct listed in the report are: '(a) funding and career pressures of the contemporary research environment (such as the pressure to publish; [...]); (b) inadequate institutional oversight; (c) inappropriate forms of collaborative arrangements between academic scientists and commercial firms; (d) inadequate training in the methods and traditions of science (e) the increasing scale and complexity of the research environment, leading to the erosion of peer review, mentorship, and educational processes

in science; and (f) the possibility that misconduct in science is an expression of a broader social pattern of deviation from traditional norms' (NAS et al., 1992, p. 30). Considering the ongoing heated discussions concerning the role of replication in science, it is intriguing to see that they also mention a combination of the lack of guidance provided by theory in some biological and clinical research and the nature of the subject of investigation preventing replication as a form of self-correction in biomedical science, as something to be considered (see e.g., Guttinger, 2020). The other serious deviation part that was in the 1989 definition was also removed in the 1999 definition of scientific misconduct which was ratified in 2000: research misconduct was defined as 'fabrication, falsification, or plagiarism in proposing, performing, or reviewing research or in reporting research results' (Office of Science and Technology Policy, 2000, p. 76262), which now seems to be the most widely circulated definition of research misconduct (Resnik et al., 2015). However, at the same time, there still is quite the diversity regarding what is defined and treated as misconduct across countries and institutes (see e.g., Dal-Ré & Marušic, 2025; Fanelli, 2011; Li & Cornelis, 2021; Resnik et al., 2015). Similar to research integrity, what counts as misconduct seems to be so variable and confused in the current literature that some already recommend defining the concept if it is used, to prevent confusion (Dal-Ré & Marušic, 2025). Furthermore, the terminology in the discourse about scientific misbehaviour or deviance has shifted in the last decades, going from fraud to scientific misconduct and research misconduct over to more 'benign concepts' (Hesselmann & Reinhart, 2024, p.4), of questionable research practices and research integrity or responsible conduct of research (Carcausto-Calla et al., 2023; Degn, 2020; Roberts et al., 2020; Zuckerman, 2020).

Research integrity and the emergence of Open Science

There is a growing research community dedicated to the investigation of research integrity and related issues (see e.g., Aubert Bonn & Pinxten, 2019; Carcausto-Calla et al., 2023; Nguyen & Tuamusk, 2024; Soehartono & Khor, 2022; Tang, 2025b; Vendé et al., 2025). This emerging community surrounding research on research integrity is still somewhat scattered (see e.g., Aubert Bonn & Pinxten, 2019; Vendé et al., 2025), but also has some characteristics of a research specialty (Morris & Van Der Veer Martens, 2008). They, for instance, have dedicated journals where most of their research is being published, like *Journal of Academic Ethics*, *Accountability in Research, Research Integrity and Peer Review,* and *Science and Engineering Ethics* (Carcausto-Calla et al., 2023; Nguyen & Tuamusk, 2024; Vendé et al., 2025). Furthermore, they organise regular international conferences dedicated to research integrity, such as, the *World Conference on Research Integrity (WCRI)* and

the *ENRIO congress*, organised by the European Network of Research Integrity Offices (see e.g., Aubert Bonn & Pinxten, 2019; Steneck et al., 2017).

However, the community engaging with research integrity is one among several somewhat recent intellectual groupings dedicated to the state and functioning of the science system. There is quite the diversity of intersecting and coexisting communities (movements) with partially overlapping and diverging agendas, approaches and foci when it comes to understanding and improving or reforming science (see e.g., Anderson-González & Gruson-Daniel, 2023; Bartscherer & Reinhart, 2025; Field, 2022; Field et al., 2024; Rushforth & Hammarfelt, 2023; Peterson & Panofosky, 2023), with research integrity being one of them. One theme that comes up in all these communities in various ways, is Open Science, and its different aspects like transparency, data sharing, open access, replication, reproducibility, preregistration, and registered reports.

Following claims about a 'replication crisis' or 'credibility crisis', movements emerged aiming to open-up research (see e.g., Nosek et al., 2012; Ottolini & Noel, 2024), to make it more accountable and credible (see e.g., Nosek et al., 2015; Peterson & Panofsky, 2023). In recent years Open Science seems to have increasingly become part of the research integrity discourse (see e.g., Tang, 2025b), with various research integrity scholars explicitly stating that Open Science, and transparency are closely associated to research integrity or are even components of it (see e.g., Bouter, 2024; Haven et al., 2022). Furthermore, in recent years research policy is incorporating Open Science. *The European Code of Conduct for Research Integrity* has been in clear alignment with and in support of the Open Science agenda in the European context for years (see e.g., ALLEA, 2023; Laine, 2018; Lund, 2025), as are other national and international guidance documents (see e.g., Labib et al., 2025; UNESCO, 2021). Although previous iterations differ from current developments, the idea of Open Science, and relating it to issues of research integrity and misconduct, of course, is not new (Chubin, 1985). Two topics and practices that increasingly come up in the recent research integrity literature are preregistration and replication.

As part of the Amsterdam Agenda on the promotion of discussion and coordination of efforts to improve research integrity on a global scale, point 1 outlines the establishment of a Registry for Research on the Responsible conduct of Research (RRRCR), with six key parts to be described: 1) the problem to be addressed, 2) the impact of that problem, 3) the planned intervention, 4) the specific hypothesis, 5) the planned hypothesis test, and 6) data sharing (World Conference on Research Integrity, 2017). The first five elements to be captured in this registry mirror components of preregistration. Hence, preregistration is seen by many as an important component of research integrity, as

scholars responsible for the World Conference on Research Integrity, guidelines, and eminent figures in the research integrity community attest. These scholars also somewhat regularly conduct studies to assess the uptake of preregistration among papers submitted to their conferences (Sijtsma et al., 2021; Emons et al., 2025). Furthermore, the UKRIO's *Code of Practice for Research* explicitly states that "where appropriate, a study should be registered with an appropriate body to align with transparency and openness of the research. For example, a researcher could use pre-registered reports" (UKRIO, 2025, p. 12). Additionally, there are metrics for research integrity based on measuring OS practices. For instance, based on article metadata PLOS uses Open Science indicators, quantifying OS practices in their journals, among them whether a publication has a preregistration attached to it, as a measure of research integrity (Cadwallader et al., 2023).

In the context of increased attention to the role of replication and reproducibility in science due to recent claims about a 'reproducibility crisis' (Baker, 2016), and calls for more replication and reproducibility as a hallmark for 'good science' by metascience and Open Science proponents alike (see e.g., Edlund et al., 2022; Simons, 2014; Zwaan et al., 2018), in the recent research integrity literature reproducibility and replication are among the promoted practices and standards aimed at securing as well as improving research integrity (see e.g., Bouter & de Riet, 2021; Diaba-Nuhoho & Amponsah-Offeh, 2021; MacLeod, 2022; Resnik & Shamoo, 2017). Irreproducibility of research has, for instance, been linked to the issue of increasing numbers of retractions. Casadevall et al. (2014), based on the analysis of 423 retraction notices of articles indexed in PubMed, report that 16.1% (68 articles) of the articles were retracted due to lack of reproducibility. Hesselmann and colleagues (2017) conducted a review on the research literature on retractions of journal articles. In that review they list existing reviews on retractions and six of them contain data on prevalence of retraction due to lack of reproducibility. The lowest percentage of retractions because of irreproducibility, as reported by Azoulay et al. (2012), is 8.3% and the highest of 20% is reported by Redman and co-authors (2008). However, Hesselmann and colleagues (2017) note that there is a disconnect in the literature on retractions and add the important caution that those reported proportions are difficult to compare due to, for instance, the use of different and opaque search strategies and time intervals. Furthermore, the label 'retraction' is a bit of a misnomer, as retracted articles are not literally taken out of the published record and can still be cited (Schmidt, 2024; Woo & Walsh, 2024), but rather just marked as 'retracted'. Therefore, retractions are rather a process of visibilisation (Hesselmann & Reinhart, 2021), similar to Open Science Badges (van Elk et al., 2018)

The concept research integrity

Attempts at an explication or codification of a value system guiding proper scientific practice that can be found in current guidelines for research integrity or responsible conduct of research, where one encounters such values and principles like reliability, respect, honesty, and accountability, are not new (ALLEA, 2023). Probably the most widely circulated historical example is Merton's (1942) functionalist appeal for science in a democratic society against the backdrop of the totalitarian social order of the spreading Nazi ideology. According to him, the ethos of science is characterised by the normative values of universalism, communism, disinterestedness and organised scepticism (Merton, 1942). Merton himself already early on hinted towards the circumstance that science in action is always a push and pull between norms and counter-norms, what he called 'the ambivalence of the scientist' (Merton, 1963, p. 433), a special form of sociological ambivalence (Merton & Barber, 1963). One somewhat more detailed example from Merton (1957) is the push and pull between humility and originality in making claims about discoveries, called 'a distinct ambivalence toward the claiming of priorities' (Merton, 1957, p. 647). However, while Merton focused on the impersonal and democratic values, mostly others explicated and popularised the existence of such counter-norms affecting or even guiding scientific practice and researchers' decision making. The most famous example is Mitroff's (1974) formulation of the counter-norms, directly reacting to Merton's most popular norms: interestedness, particularism, solitariness, and organised dogmatism. Both norms and counter-norms are continuously the target of endeavours of modernisation (see e.g., Anderson et al., 2010; Hosseini et al., 2024; Macfarlane, 2024).

Furthermore, Zuckerman (1984), already in 1984, warned about overly global formulations of what constitutes good scientific conduct or a scientist acting in accordance with the norms of science (also see Zuckerman & Sagarin, 1977). Criticising the reduction of 'good science' to demonstrating honesty as a scientist, she wrote that what constitutes honest conduct 'has different meanings in different concrete circumstances [...] Therefore, a global formulation obscures rather than clarifies the workings of norms' (Zuckerman, 1984; p. 9). Now 40 years later this obscuring formulation of the honest scholar is still supposed to guide research conduct (ALLEA, 2023; Desmond & Dierickx, 2021; Steneck, 2011). Moreover, already early on, these exercises in explicating values of 'good science' have received methodological and epistemological critique. One such early critique stems from Mulkay (1976), who argued that these attempts at explication or codification are not capturing the normative structure of science, but rather an ideology of science that has less to do with what really happens in the day-to-day practice in the lab and is more

the result of scientists presenting and preserving a certain image of science to the outside (Gieryn, 1983; Mulkay, 1976).

Harry Collins describes science as craftwork with integrity (Collins, 2021; Collins & Mills, 2022). However, it remains unclear what the term integrity actually entails in the context of research and science. In line with the historical and institutional origin, often the approach to address and frame research integrity, especially in policy, seems backwards, with contemplations about what constitutes a lack or violations of 'good science' (see e.g., Horbach & Halffman, 2017; Horbach et al., 2024). Accordingly, questionable or detrimental research practices are often interpreted as resulting from a lack of research integrity, which may be due to the people involved, so-called 'bad apples' and biased human nature, or the science system and environment itself, due to so-called 'perverse incentives' (see e.g., see e.g., Bouter, 2015; Degn, 2020; Downey & Veitch, 2021; Flis, 2019; Kennedy et al., 2023; Lambert & Degn, 2026; Morawski, 2020; Stephan, 2012; Vendé et al., 2025)[2]. Both of these potential causes for problems in the sciences are contested and hotly debated. For instance, whether one can just blame the incentives or so-called external factors remains questionable (see e.g., Schneider et al., 2021; Yarkoni, 2018).

Steneck (2006) attempted to provide some clarity to the terminology and how the concept research integrity relates to associated concepts like questionable research practices, misconduct, research ethics, and responsible conduct of research. The ideal of responsible conduct is here the result of research ethics and research integrity. Hence, according to Steneck (2006), research integrity is part of the broader concept of responsible conduct of research (Armond et al., 2024). Accordingly, research ethics comprises 'the moral problems associated with or that arise in the course of pursuing research' (Steneck, 2006, p. 56), and research integrity is defined as 'possessing and steadfastly adhering to professional standards as outlined by professional organizations, research institutions, and when relevant, the government and public' (Steneck, 2006, p.56).

Moreover, Shaw (2019), who, in reference to Steneck (2006), stumbles again over this 'ironic' lack of clarity concerning the terminology despite the long history of the research integrity discourse, has attempted to remedy and clarify (Shaw, 2019, p. 1086). Shaw (2019) identifies some key issues and distinctions surrounding the meaning of research integrity. First, according to Shaw (2019), scientific and research integrity are conflated within the literature, while being actually quite different. Research integrity is seen as a

[2] This 'bad apple' terminology is borrowed from organisational theory (Kisch-Gephart et al., 2010).

combination of institutional and personal integrity of the researcher and scientific integrity as the integrity of the knowledge produced. Hence, research integrity, according to Shaw's (2019) distinction, seems to be about the processes of doing science, while scientific integrity is more about the products of science. He further, similar to Steneck (2006), goes into detail regarding the distinction between research ethics and research integrity, by claiming that research ethics systems are set up to protect participants and ensure that they are treated properly, while research integrity is more about what goes on once the research project has been initiated. The difference between research ethics, research integrity and scientific integrity is illuminated based on the example of plagiarism. According to Shaw (2019), plagiarism might infringe on research ethics and research integrity, but it does not affect scientific integrity, because the scientific product itself remains the same even if it is stolen. In the distinction between research and researcher integrity, he further outlines how the responsibility of the researcher has to be understood as situated within a larger system of, for instance, journals and institutes (Shaw, 2019).

Shaw (2019) also goes into the issue of redundancy in doing science in the context of the question of what the added value of the research actually is, when the specific research question has already been answered. This, of course, is especially intriguing in the current context of calls for more replications in science (see e.g., Buck, 2024; Derksen et al., 2025; Edlund et al., 2022; Simons, 2014), because it suggests a more complex relationship between research integrity and replication, than just that more replication leads to more research or scientific integrity. While there are these attempts at conceptual clarification, unfortunately, within the current research integrity discourse it largely remains unclear what research integrity actually means and how that concept relates to related concepts like responsible conduct of research, research ethics, research governance, research quality, Open Science, transparency or reproducibility (see e.g., Haven et al., 2022; Helgesson & Bülow, 2023; Kolstoe & Pugh, 2024; Shaw, 2019; Tang, 2025b). Furthermore, it often remains unclear what is actually talked about when the term research integrity is used: the integrity of research, the integrity of researchers, or the integrity of the research system? They seem to be quite conflated in the literature (Helgesson & Bülow, 2023). Helgesson and Bülow (2023) suggest that these three understandings of research integrity might be separated as follows: the integrity of research is largely about research quality, the integrity of the research system is largely about research governance, and the integrity of the researcher is about the individuals' (groups') character and values (Helgesson & Bülow, 2023). These three subjects of integrity are distinct, but, of course, highly interdependent. Importantly, research integrity seems to be more complex and nuanced than how it is often conceptualised inversely as just the

opposite or prevention of misconduct in the form of falsification, fabrication and plagiarism or questionable research practices, like p-hacking and HARKing (see e.g., Downey & Veitch, 2021; Horbach & Halffman, 2017; Neoh et al., 2023; Resnik, 2019). Furthermore, in codes of conduct, it is often described as a combination of somewhat vague principles like honesty, reliability, accountability or respect and attempts at explicating general responsibilities during different research processes (see e.g., ALLEA, 2023; Desmond & Dierickx, 2021). Other approaches are, for example, to list or rank aspects and processes of science that are relevant for the integrity of research (see e.g., Labib et al., 2021; Sørensen et al., 2021). Therefore, overall, the meaning of research integrity is fundamentally confused and remains contested (see e.g., Horbach & Halffman, 2017; Roberts et al., 2020; Shaw, 2019). As Field and colleagues (2024) put it, based on a review of literature on responsible conduct of research: 'its treatment as an umbrella concept welcoming various responsibilities, virtues, structures and responsible practices has rendered it so diffuse a concept that it risks becoming meaningless' (p. 9).

Due to many of the problems at the centre of recent science reforms and concerns around research integrity originating from a small number of disciplines, there seems to be a central role being played by a notion of science stemming from biomedical as well as some social and behavioural sciences (see e.g., Clancy et al., 2026; Nelson et al., 2021, 2022). This might be problematic in the context of understanding research integrity or what 'good science' is (Penders et al., 2009), because research itself is a multifaceted and diverse landscape of different combinations of values, practices and conditions (Knor-Cetina, 1999), sometimes referred to as the disunity of science (Galison & Stump, 1996). Laudel (2024), for instance, argues that the diverse epistemic properties, characterising research fields or communities across the research landscape, affect or even determine specific notions of research quality, resulting in quite different strategies and practices to establish and assess the quality of the research. Correspondingly, Sørensen et al. (2021) found a difference in prioritisation between the humanities, the social sciences, the medical sciences, and the natural sciences regarding topics they associated with research integrity. Likewise, another problem inherent to the prominent tendency to define research integrity as the opposite of problematic conduct is that what counts as questionable research practices seems to be a grey area with differences depending on the characteristics of the research at hand (Ravn & Sørensen, 2021; Reischer & Cowan, 2020).

Considering the historical origin of the concept of research integrity from a policy, regulatory, and bureaucratic context as a reaction to prominent or even scandalous research misconduct cases and its current implementation in the global science system, it seems appropriate to refer to research integrity as

a ‘policy object’ (Davies & Lindvig, 2021, p. 444) and in combination with misconduct as ‘technologies of governance’ (Davies, 2019, p. 1237). Davies and Lindvig (2021) state that as a policy object (Sin, 2014), research integrity has an enacted ontology. As such, research integrity cannot exist independently from specific practices and beliefs situated in varying contexts and roles in the science system. Therefore, a specific multiplicity and diversity concerning what constitutes research integrity across research communities, nations, stakeholders, institutes and just the whole scientific ecosystem seems inevitable (Davies & Lindvig, 2021; Hastings et al., 2023; Penders et al., 2009). However, there is a tendency of policy discourse and the research on research integrity literature to capture what constitutes research integrity or ‘good science’ in an essentialist, ahistorical, uniform manner (Horbach & Halffman, 2017). However, while this practically hides the reality of coexisting forms of ‘good science’, this seems to also be in contrast to and at times in confrontation with the researchers’ own experiences and perceptions of their research realities (Davies, 2019; Davies & Lindvig, 2021; Hangel & Schickore, 2017; Hastings et al., 2023; Horbach & Halffman, 2017; Langfeldt et al., 2020; Penders et al., 2009; Pols et al., 2024; Schickore & Hangel, 2019; Vincent & Hooper, 2025). Furthermore, in the institutional policy context, research integrity is not outlined as what constitutes ‘good science’; it is outlined in a rather punitive way, listing what needs to be punished (see Horbach & Halffman, 2017). This might be linked to an overarching institutional focus on accountability instead of responsibility (Douglas, 2024). Indicating that in, contrast to what others have suggested (Davies & Lindvig, 2021; Felt, 2017), the codification of ‘good science’ under the umbrella of research integrity is less an attempt at ‘responsibilization’ (Davies & Lindvig, 2021, p. 445), but rather a process of making researchers accountable for their conduct (see e.g., Horbach et al., 2024). Such a focus on accountability fits, however, well into the bureaucratic audit culture of modern managerial academia with a strict focus on accountability in a system of academic capitalism (Halffman & Radder, 2015; Hostler, 2024b; Slaughter & Leslie, 1997; Strathern, 2000). Vincent and Hooper (2025) use the distinction between a thick ethos, as a coherent internalised complex system of culturally specific values, knowledge, heuristics, as well as skills, and thin values, which are non-internalised rules, incentives, and norms, of research integrity. Vincent and Hooper (2025) further conceptualise the codification of research integrity and what constitutes ‘good science’, in for instance conduct guidelines, in the context of a deontic shift, as a process of thinning largely propelled by the globalisation and growth of the science system (De Sola Price, 1963; Douglas-Jones & Wright, 2017; Hanson et al., 2024). Furthermore, existing research suggests that guidelines and protocols of proper scientific conduct are perceived as rather burdensome, inapplicable,

irrelevant, or an infringement of academic freedom or even an insult to the scientists' competence to produce trustworthy science (see e.g., Davies, 2019; De Vries et al., 2006; Hangel & Schickore, 2017). Respondents in De Vries and colleagues (2006) study even suggest that rules intended to guide proper conduct of scientists can also 'generate misconduct' (p. 45), because scientists in their daily practice need to find ways to work around such regulation to be able to get things done.

Researchers' perspective on research integrity

The literature indicates a perceived disconnect or misalignment between research evaluation regimes and the realities as well as understandings of 'good research' from within research communities (Benedictus et al., 2016; Biagioli & Lippman, 2020; Cramer & Reinhart, 2026; Edwards & Roy, 2017; Labib et al., 2024; Robinson-Garcia et al., 2023; Pols et al., 2024; Valkenburg et al., 2021). In the context of Open Science as a growing part of research integrity discourse, there is also a counter-discourse emerging about potentially detrimental disconnects and misalignments with diverse and local epistemic communities, because incentivised or required Open Science practices and standards are seen as unfeasible or plain irrelevant (see e.g., Bazzoli, 2022; Guttinger, 2020; Hostler, 2023; Leonelli, 2018, 2022, 2023, Penders et al., 2019, 2020; Rubin, 2023; Ulpts & Schneider, 2025). While research on experiences and perspectives of research integrity and 'good research' from within specific research communities is still quite scarce, there are some noteworthy exceptions (see e.g., Brenninkmeijer & Pols, 2026; Davies, 2019; Davies & Lindvig, 2021; De Vries et al., 2006; Hangel & Schickore, 2017; Pols et al., 2024; Schickore & Hangel, 2019). While the literature suggests researchers tend to provide formulaic and sanitised responses in accordance with an ideology of science (Mulkay, 1976), emerging themes are, for instance, replication and replicability being a hallmark of good science and even a methodological imperative, or about the necessity for transparency of research processes and the accessibility of data. However, they also talk about the complexities, uncertainties, and complicated nature of such issues. Furthermore, research integrity and research quality are not just conceptually close, but recently there have also been investigations into bottom-up notions of research quality in reaction to a perceived misalignment between internal preferences for research quality and those used in research evaluation (Gläser, 2024; Laudel, 2024). Some of these studies are also often framed in terms of 'good research' (see e.g., Hug et al., 2013; Langfeldt et al., 2020; Ochsner, 2022). In the following I will outline some of the research on research integrity from within research communities in more detail to provide a picture of the existing literature.

In their focus group study, De Vries and co-authors (2006) highlight four key themes as emerging from the researchers: The meaning of data, the life with colleagues, the pressures of production in science, and the rules of science. Dealing with data is described as somewhat of a grey area, where researchers seem too busy and struggling with finding the line between cleaning data and cooking the data, while it further remains unclear how to report all the decisions made about handling the data and how such decisions affect the results (De Vries et al., 2006). Maybe somewhat counterintuitively, rules are presented as a cause of misconduct and QRPs, because a plethora of as unnecessary and intrusive perceived rules require researchers to find ways of getting around those rules to actually be able to do the research (e.g., complaining about the demands of the institutional review boards) (De Vries et al., 2006). Scholars describe a constant need to negotiate the relationships and distribution of responsibilities between colleagues, and that 'good science' further depends on actively attending to those relationships, because competition may harm those relationships (DeVries et al., 2006). Relatedly, the pressure to produce in science is outlined in the context of stories of exploitation and worries about being scooped. In reference to Durkheim (1982) De Vries and co-authors (2006) claim that 'a certain amount of 'normal misbehavior' is common in the dynamic field of science' and that they, in reference to Merton's (1963) notion of the 'ambivalence of the scientist', play 'a useful and irreplaceable role' (De Vries et al., 2006, pp. 47-48).

While respondents presented a diverse picture of interpretations of research integrity, based on interviews with life scientists in Switzerland Shaw and Satalkar (2018) identified three main themes capturing research integrity according to the researchers. The first focuses on values or principles that define research integrity. Those are: honesty in reporting, to colleagues and in authorship, objectivity by following the data, admitting when hypotheses are wrong, protecting against human bias and keeping observations pure, and starting with as well as sticking to a relevant research question. Other aspects defining research integrity provided by the interviewees were the following: the importance of following rules and protocols, and the importance of the personal integrity of the researcher (Shaw & Satalkar, 2018). The second is the distinction between scientific and research integrity. They are either seen as referring to the same thing, or research integrity is viewed as a subset of the wider integrity of science. Lastly, the third is the difference between integrity and misconduct. Here, only a few said that the absence of misconduct indicates research integrity, while the majority mentioned that integrity can also be threatened without actual misconduct, by, for instance, unintentional or accidental errors (Shaw & Satalkar, 2018).

Hangel and Schickore (2017) use interview data to capture researchers' conceptions of good research practices. They focus on three topics: replication, pilot studies, and protocols (rules). Replications, in line with 'sanitized' and 'formulaic' depictions of an ideology of science (Mulkay, 1976), were mostly described as fundamental to science, but simultaneously also as a practice that is in conflict with career goals and perceived publication pressures, as well as being plain boring (Hangel & Schickore, 2017, p. 770). Furthermore, similar to Peterson and Panofsky (2021), in addition to a problem of (task) uncertainty that is associated with the evaluation and conduct of replication, the authors identify a logic of investment, where gains and costs of conducting a replication are weighed against each other, resulting in replication often being perceived as too expensive for the expected pay-off, for instance in terms of career advantages (Hangel & Schickore, 2017). Pilot studies are depicted as practices to secure external funding and ensure the reliability and proper working of procedures in the published study. Similar to replications, pilot studies are also characterised by a certain degree of uncertainty concerning how to conduct and evaluate them, as well as whether to report them (Hangel & Schickore, 2017). Researchers also reported some kind of suspicion towards rich labs being able to hunt for significant and meaningful findings by just doing more pilots (Hangel & Schickore, 2017). In contrast to Shaw and Satalkar, (2018), but similar to De Vries and colleagues (2006), some scientists in Hangel and Schickore's (2017) interviews also mention annoyance over protocols and rules as well as the aspect of having to deviate from institutional protocols and guidelines for scientific conduct, if they want to make progress and do 'good research'. Therefore, scientists voiced the sentiment that it should be at their discretion how to conduct their research, because they have the required expertise and integrity to ensure proper conduct (Hangel & Schickore, 2017). Schickore and Hangel (2019) further introduce a reproducibility-related file-drawer effect. Accordingly, easy experiments are chosen and more complicated, uncertain ones are avoided, because reproducibility as an internal lab standard leads to a new form of file-drawer problem and ultimately a publication bias. These uncertainties regarding the conduct and appraisal of practices that are perceived to be fundamental to 'good science', like replication and pilots, are described as a sort of messiness of the day-to-day practice of science (Schickore & Hangel, 2019). However, based on the responses Schickore and Hangel (2019) conclude that in the face of uncertainty, researchers seem to follow a strategy of avoidance instead of resolution, as illustrated by the reproducibility-related file drawer effect.

A reoccurring theme in these qualitative studies inquiring into the researchers' opinions and perceptions of research integrity is the inherent messiness, complexity, uncertainty, nuanced and contingent nature of everyday

research demanding flexibility and adaptability (see e.g., Davies, 2019; Davies & Lindvig, 2021; Hangel & Schickore, 2017; Pols et al., 2024; Schickore & Hangel, 2019; van Drimmelen et al., 2024). Since day-to-day research is characterised by varying degrees of uncertainty and messiness it becomes quite fluid, subtle, and hard to draw the line between, for instance, cheating, sloppiness, fraud and errors as well as confirmatory or exploratory components of the research (see e.g., Davies, 2019; Schickore & Hangel, 2019; Zuckerman, 1984). Therefore, it is seen as a constant process of negotiation what good research practice looks like or means and researchers actually struggle to articulate what research integrity means for them as well as what constitutes good practice or proper science (see e.g., Hangel & Schickore, 2017; Schickore & Hangel, 2019).

Expanding on the issue of uncertainty, based on focus group interview data Elizondo and Kaltenbrunner (2024) relate issues of research integrity to the inherent uncertainty of research activities by conceptualising questionable research practices and misconduct as reactions to social and epistemic uncertainties (Sigl, 2016). Social uncertainty results from human nature and social organisation of scientific work, and epistemic uncertainty is the inherent underdetermination of scientific practice (Elizondo & Kaltenbrunner, 2024; Sigl, 2016). Followingly, they identify five main strategies to deal with uncertainties: 1) cutting corners to alleviate time pressure, 2) using data practices in the grey zone to publish clean results, 3) treating research ethics as just another box to tick, 4) avoiding trouble of addressing potentially problematic practice in one's environment, and 5) strategic authorship and grant writing practices to secure publication and funding (Elizondo & Kaltenbrunner, 2024). Hence, most of the practices that are seen as questionable or even as misconduct are conceptualised as strategies to evade social uncertainties and pressures, stemming from the organisation of science funding, rather than epistemic uncertainties (Davies, 2019; Elizondo & Kaltenbrunner, 2024; Pols et al., 2024).

The existing literature attempting to capture the researchers' perception and understanding of what constitutes 'good research' bottom-up suggest that there are tensions when it comes to notions of research integrity between actual researchers and those wanting to regulate science, like policymakers and research integrity scholars alike (see e.g., Brenninkmeijer & Pols, 2026; Davies, 2019; Davies & Lindvig, 2021; De Vries et al., 2006; Hangel & Schickore, 2017; Pols et al., 2024; Schickore & Hangel, 2019). As Pols and co-authors (2024) based on their inquiry in the Dutch context state: 'The "view from within" provided by our analysis provides detailed understandings of academic practice, where policy interferes by enforcing a reality of excellent (or fraudulent) individual scientists in competition with one another and organized by ideas about academia functioning as a market …' (p. 44). Followingly,

Brenninkmeijer and Pols (2026), based on their ethnography in Dutch chemistry labs, conceptualise ‘good research’ as a form of caring characterised by socio-material collaboration between peers and chemicals, instead of individualised competition.

Chapter 3: Theoretical framework: the transepistemic arena of 'good research'

In this chapter I present the theoretical framework of the dissertation. I build the framework by situating research integrity within the cycle of credibility in the larger transepistemic arena of research (see e.g., Hessels et al., 2019; Latour & Woolgar, 1979; Knorr-Cetina, 1982, 1983). I conceptualise research integrity as a policy object that mainly enters the research communities in spaces of research evaluation, like in the interaction with funders and institutes, and provides a reference system for the researchers with their management of credibility in interactions with different stakeholders (Davies & Lindvig, 2021; Shapin, 1995a; Sin, 2014; Smith, 2014). The attainment and maintenance of credibility is achieved through, what I call credibility work, which are the practices that aid the researchers in demonstrating and persuading their research community and other actors that they as well as their research are credible (Daston, 1995; Geiger, 2021; Meskus et al., 2018; Shapin, 1984, 1995a; Shapin & Schaffer, 1985).

Literature in STS and the social studies of science suggest that discussions about reliability, validity, but also research quality and integrity are either fundamentally about issues of credibility or at least can be conceptualised as such for the purpose of investigating and comprehending scientific knowledge production (see e.g., Aradau & Huysmans, 2019; Geiger, 2021; Latour & Woolgar, 1979; Li Vigni et al., 2023; Meskus et al., 2018; Shapin, 1995a; Shapin & Schaffer, 1985). Furthermore, large portions of the discourse revolving around research integrity and proper conduct are explicitly framed in terms of credibility, whether the conversation is about a 'credibility crisis', 'credibility revolution' or public trust in science (see e.g., Angrist & Pischke, 2010; Vazire, 2018). Therefore, to get closer to what Valkenburg and colleagues (2021) called a 'cultural-practice perspective' (p.1) to research integrity and gain an emic understanding of research integrity enacted as 'good research' (see e.g., Pols et al., 2024), I employ social studies of science work on credibility located within the transepistemic arena of the sciences in interaction with different stakeholders (see e.g., Knorr-Cetina, 1982, 1983; Li Vigni et al., 2023). The social studies of science, and especially the sociology of scientific knowledge (Shapin, 1995b), have a long tradition of investigating the social construction of matters of facts as issues of credibility and persuasion within specific (research) communities;

not independent of epistemic considerations or purely social, but in constant co-construction, in the local processes of (scientific) knowledge production (see e.g., Aradau & Huysmans, 2019; Bourdieu, 1975; Collins, 1975, 1985; Geiger, 2021; Gläser, 2024; Jasanoff, 1996; Knorr-Cetina, 1981, 1983, 1995, 1999; Latour & Woolgar, 1979; Laudel, 2024; Li Vigni et al., 2023; Lynch, 1985; Meskus et al., 2018; Shapin, 1984, 1995a; Shapin & Schaffer, 1985; Whitley, 1983, 1984).

Knorr-Cetina (1982) criticised classical sociological concepts like research specialty and research community for their reductive and etic nature, claiming their 'irrelevance' and 'meaninglessness' (p.117), because such concepts, for instance, tend to neglect the critical involvement of non-academic actors and do not capture researchers' perceptions (of belonging). Therefore, she introduces the notion of a 'transepistemic arena' which aims to situate scientific knowledge production within (emic) local sites of knowledge production (Knorr-Cetina, 1983, 1995). Researchers and research practice are not just dependent on other scholars (Whitley, 1984), but also on different stakeholders within the science system, like funders, industry, publishers, politics and other policymakers (see e.g., Knorr-Cetina, 1982, 1983; Latour & Woolgar, 1979; Penders & Neils, 2011; Smith, 2014). As such the notion of a transepistemic arena aids us in narrowing our focus on the local site of knowledge production, while at the same time broadening our perspective by incorporating various types of actors that researchers continuously interact with and depend on (Knorr-Cetina, 1982). Furthermore, what constitutes 'science' is continuously renegotiated and the (symbolic) boundaries between academic and non-academic actors are constantly moving or being pushed back and forth (Felt & Frantz, 2026; Gieryn, 1983, 1995, 1999; Sismondo, 2025). Thus, what counts as internal or external to science has always been a moving target. Even in the case of funders, journals and research institutions, the research(er) assessment or evaluation in the form of peer review is usually done by or at least involves researchers (Lamont, 2009; Schneider et al., 2021). Hence, science has always been a transepistemic arena (Knorr-Cetina, 1982), comprised of academic and non-academic actors constantly negotiating the nature and role of science in society. For instance, following high profile research misconduct cases regulatory bodies like governments, funders and research institutes as well as universities initiated new regulatory procedures in science around the globe (see e.g., Horbach et al., 2024; Steneck, 2011). One of such regulatory interventions is the policy object research integrity (Davies & Lindvig, 2021). Historically the scientific community largely revolted and argued against its introduction and perceived it as an intrusion of politics, policy and more generally other non- or extra-scientific regulatory bodies that attempt to control or steer their conduct in the name of 'good', 'better' or 'legitimate science'.

However, responses were not uniform, there were also scholars who welcomed such interventions and recognised value in addressing the issues surrounding misconduct and a perceived lack of control and regulation in science (see e.g., Felt & Frantz, 2026; Francis, 1989; Guston, 2000; Hesselmann & Reinhart, 2024; Steneck, 1994; Tamot et al., 2013). This can be treated as an act of boundary work (Gieryn, 1983, 1995, 1999), where scientists attempted to ensure and defend their autonomy, against interventions in the self-regulation of science in matters of self-correction and research integrity (see Chapter 2 on the history of research integrity). However, such political and policy interventions in reaction to widely publicised episodes that put into doubt the ideals of science (Guston, 2000; Hesselmann & Reinhart, 2024; Price, 2013), make the ongoing processes of boundary work more salient (Felt & Frantz, 2026; Gieryn, 1983, 1995, 1999; Guston, 2000).

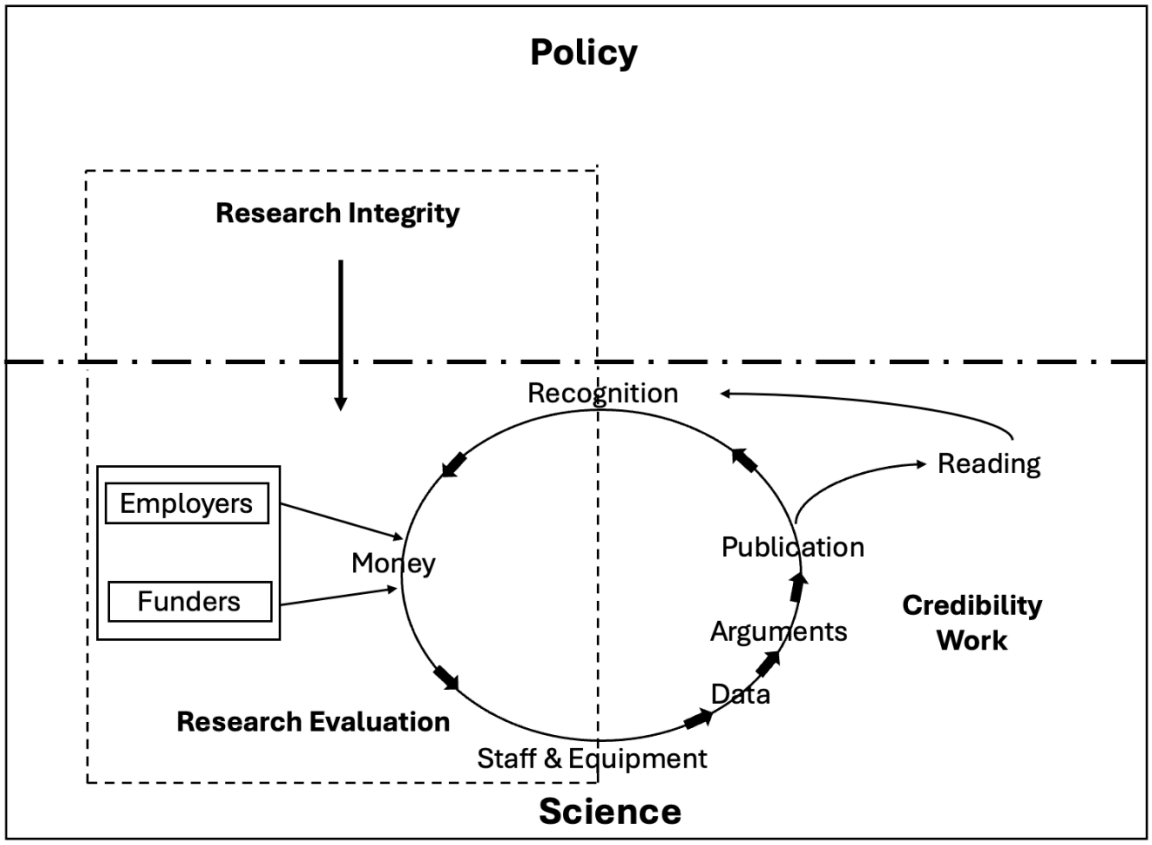
Following the intervention of the US congress in the (assumed) self-regulation of science and taking the (symbolic) social contract for science as a heuristic starting point (Guston, 1994), Guston (2000) uses principal-agent theory as a tool to structure the investigation of science policy and frame the relationship between science and politics in the US with the Office of Research Integrity (ORI) located on the boundary between science and politics. According to the (heuristic) social contract for science, science is supposed to provide certified knowledge and feed into innovation as well as inform public policy and thereby contribute to society, in return science is funded and largely granted autonomy without interferences from politics when it comes to scientific (self-)regulation and (self-)correction (Douglas, 2024). For Guston (2000) this idealised heuristic contract provides a convenient reference point for the moveable social (cultural) boundary between science and policy for the investigation of the boundary work and re-negotiation of the scientific self in the context of 'good science' (Felt & Frantz, 2026; Gieryn,1983, 1995, 1999; Sismondo, 2025). According to Guston (2000) principal-agent theory alone is too essentialist (etic) in allocating roles and understanding how boundaries are set and moved. To that end, inspired by Gieryn (1995) and based on his constructivist approach, Guston (2000) aims to overcome the inherently essentialist nature of principal-agent theory, by conceptualising ORI as a boundary organisation which produces boundary objects (Star & Griesemer, 1989), and standardised packages (Fujimura, 1992) for research misconduct investigations and the protection of whistleblowers (Guston, 2000). In reference to Latour's (1987) notion of the Janus face and the specific pattern of communication as well as conduct depending on the target audience (Shapin, 1995a), Guston describes the boundary organisations, like ORI, as sites of co-production (Jasanoff, 1996) located on the boundary between science and politics. I

use Guston's (2000) approach to science policy for research integrity as a point of departure.

While in the US context of the late 1990s and early 2000s Guston's (2000) conceptualisation of ORI as a boundary organisation might have been appropriate within the narrow confines of misconduct allegations and investigations, boundary objects in the context of research integrity and standardised packages remain doubtful in the current and broader context. The 'umbrella concept' research integrity is not anymore just the inverse of misconduct or synonymous with the site of misconduct allegations, but it encompasses a vast array of practices, ideals, norms, values and prescriptions in the name of 'good science' as well as the prevention of 'bad science' (Valkenburg et al., 2021, p. 2; also see e.g., Horbach & Halffman, 2017). While research integrity or 'good research' possesses the interpretative flexibility, of a boundary object, as it travels across the research landscape, boundary objects require that the object in question has a specific core that remains stable across domains and that there is some limitation concerning the scope and scale (Star, 2010). However, this is not provided with research integrity more generally without restricting it to a specific site of interaction between different social worlds dedicated to a certain purpose, like ORI (Guston, 2000), or the Museum of Vertebrate Zoology at the University of California, Berkeley (Star & Griesemer, 1989). Research integrity is characterised by multiplicity (see Mol & Law, 1994), because there are vast confusion and diversity what research integrity or misconduct and questionable research practices are across the research landscape, institutions, disciplines and nations (see e.g., Dal-Ré & Marušic, 2025; Degn, 2020; Field et al., 2024; Hastings et al., 2023; Hesselmann et al., 2014; Horbach & Halffman, 2017; Kolstoe & Pugh, 2024; Penders et al., 2009; Ravn & Sørensen, 2021; Shaw, 2019; Steneck, 2006). However, the more general notion of boundary work and ongoing negotiation of cognitive territory in the cultural map of science from Gieryn (1983, 1995, 1999) remains applicable and a useful theoretical lens for investigating the interplay between science and the science policy landscape surrounding research integrity and what counts as legitimately scientific from actors perceived as inside as well as outside the research communities (Felt & Frantz, 2026; Sismondo, 2025). In accordance with its historical origin from policy or even politics, following Davies and Lindvig (2021), I conceptualise research integrity as a policy object that can be enacted in various ways depending on the purposes and local contexts it is situated in (see Sin, 2014). According to Sin (2014) policy objects acquire their multiple meanings and ontologies through their enactments depending on the interpretations and negotiations within specific local contexts. Therefore, while the policy object of research integrity might have a specific ontology and meaning according to ALLEA, or national policy documents, such ontologies

and meanings change and adapt in local contexts through processes of enactment (Degn, 2020 called this translation in the institutional context). In the science system the policy object research integrity and related concepts like responsible conduct of research (RCR) largely work through regulatory bodies, like research institutes and funders (Davies & Lindvig, 2021; Sin, 2014), where processes of research and researcher evaluation take place (Cramer & Reinhart, 2026; Kulczycki, 2023). See Figure 1 for a depiction of research integrity in the transepistemic arena.

Figure 1. The transepistemic arena of research integrity



Credibility cycle

Although Guston (2000) mainly uses the cycle of credit to illustrate the self-regulatory nature of science, I expand on his application. To answer the question of what motivates scientists, Latour and Woolgar (1979) came up with the cycle of credit (henceforth credibility cycle), which frames scientific work as the cyclical successive process of conversion (translation) of one type of capital into another, which can then be re-invested to further scientific work or one's career. Similar to other sociology of science scholars who conceptualise science as market economies or embedded in them (Hagstrom, 1965). As Bourdieu (1975) who describes scientific conduct as the struggle for capital and

scientific monopoly in the scientific field, Latour and Woolgar (1979) try to understand the scientists they investigated as investors of credibility based on the quasi-economic metaphors and analogies offered by the researchers they encountered in their lab ethnography (or as Shapin, 1995a put it, researchers weigh credibility risks and credibility rewards). Based on their appraisal of their lab observations, document analysis, and interviews, Latour and Woolgar (1979) deem credit as reward insufficient to explain scientists' behaviours and motivations as well, as it only provides a foundation for retrospective perspectives. Therefore, they conceptualise credit as forms of credibility in science. The quasi-economic analogies or metaphors used by the researchers in Latour and Woolgar's (1979) lab ethnography do not necessarily mean that their conduct and psychology one-to-one follow such economic investment strategies, but it could also be a consequence of the researchers' attempt to make the usually tacit explicit and their behaviour and motivation understandable to outsiders (Gläser, 2024). In its various forms credibility is a prerequisite to be able to conduct research (Latour & Woolgar, 1979; also see Whitley's, 1984 conceptualisation of reputation). Accordingly, science is the continuous process of investing and re-investing forms of credibility from data, arguments and articles to recognition after reading to acquiring money in the form of grants which allows for buying of equipment, all enabled by acquiring and maintaining credibility. Here success means accelerating the act of going through the cycle (see Figure 1) (Hessels et al., 2019; Latour & Woolgar, 1979).

In science studies the credibility cycle has been applied and adapted to many different contexts it, therefore, exists in various iterations in the literature. To list just some of them: Rip (1994) used the credibility cycle to locate and understand the role of research councils between the state and national research communities, Packer and Webster (1996) adapt the credibility cycle to integrate the patenting culture at universities into the academic credibility translation processes, Wouters (1997) integrates Price's notion of the citation cycle into the credibility cycle, and Penders and Neils (2011) apply and adapt the credibility cycle to comprehend credibility engineering in food industry science as two intersecting cycles of researchers and consumers (also see Hessels et al., 2009). In their investigation of the interrelations between health inequality research and policy Smith (2014) adapted the credibility cycle to address the multiple audiences that the researchers perceived to interact with and depend on. Smith' s (2014) adaptation explicitly adds the role of policymakers and funders to the already included academic peers, when it comes to researchers securing credit in the form of recognition and resources through the acquisition and maintenance of credibility in the interaction with these three different audiences (Shapin, 1995a). For instance, Smith notes that

when writing funding applications researchers 'try to assess what research funders are "looking for"' and due to the circular nature of the credibility cycle dominant policy ideas might affect which research and outcomes they are willing to present, because it might affect their future credibility when interacting with policymakers (Smith, 2014, p. 567). Additionally, Hessels and colleagues (2019) propose a 'refinement' to Latour and Woolgar's credibility cycle by making it 'sensitive' (p.130) to the social and epistemic diversity across the research landscape (see e.g., Galison & Stump, 1996; Knorr-Cetina, 1999; Whitley, 1984) and reacting to the well-known criticisms of its quasi-economic nature by Knorr-Cetina (1982) and Whitley (1983). They conclude that in different epistemic cultures (Knorr-Cetina, 1999), sometimes steps in the cycle can be skipped and the speed (acceleration) of the cycle varies. As such the attainment and management of credibility by conducting and presenting research in a way that is perceived as good, according to the different audiences that researchers depend on, is tantamount for the ability to conduct research (Latour & Woolgar, 1979). Therefore, in interaction with stakeholders like funders and institutes confrontations and negotiations take place between official notions of research integrity and the researchers' internal notions as well as conditions for 'good' knowledge production.

Credibility work

Geiger (2021) judges the advocacy-objectivity distinction for assessing the legitimacy of public sociology as inappropriate and 'scientistic' (p. 3). Therefore, inspired by STS work from, for example, from Shapin (1995a) and Gieryn (1983, 1999), he attempts to comprehend legitimising acts that sociologists use in communication with non-academic audiences as 'credibility work' (Geiger, 2021, p. 5). I take over this term but deviate in conceptualisation. I conceptualise credibility work as the practices that aid the researchers in the attainment and maintenance of credibility within a specific research community. Akin to a Feyerabendian 'anything goes' (Feyerabend, 1976) and Collins' methodological relativism (Collins, 1981; Franklin & Collins, 2016), Shapin (1995a) advocates something like an agnostic stance as a methodological principle, when it comes to the investigation of what researchers utilise for the appraisal and attainment of credibility in their epistemic culture (Knorr-Cetina, 1999; also see Gläser (2024) for framing practices of research quality as those that are ascribed a specific utility for the research and the researcher). In reference to Douglas (1975), Shapin (1995a) states that what counts as credible is simultaneously a question of morality and is interdependent with the (epistemic) culture the research is situated in (Knorr-Cetina, 1999). Followingly, due to the variation in epistemic cultures, the epistemic diversity across the research landscape and the general prevailing disunity of the sciences (Galison

& Stump, 1996; Knorr-Cetina, 1999; Laudel, 2024), there cannot be one universal theory of how to establish credibility irrespective of the context, conditions and aims of the case in question (Shapin, 1995a).

Based on their socio-historical comparison of Hobbes and Boyles philosophy of science, Shapin and Schaffer (1985) describe the production of matters of facts as the multiplication of witnesses, which is aided by three technologies (tools of knowledge production). These technologies are: 1) material technologies which entail all the practices, techniques, operations and resources in the production of scientific knowledge (outcomes), 2) social technologies which comprise the traditions and conventions that determine the interactions between scientists and the appraisal of knowledge claims, and 3) literary technologies which are the practices and tools that aid the researchers in making the operations and conduct understandable and traceable for anyone who did not directly witness a study. Hence, literary technologies are all the practices and tools that enable virtual witnessing (Shapin, 1984). In the transepistemic arena credibility work then is located at all the points of the credibility cycle where researchers interact with different audiences and notions of 'good science' as well as credibility are negotiated (see Figure 1). Stakeholders like, funders, institutes and peers can be seen as vectors of credibility along which researchers work to manage their credibility (Shapin, 1995a).

Similar to what Jasanoff (1996) describes as the 'co-production' of social, political, and epistemic order in science, Shapin (1995a) notes that moral ordering principles are inherent to credibility work. Hence, social, moral and epistemic order are co-produced through practices and discourse concerning credibility (Jasanoff, 1996; also see Pickersgill, 2012, 2021). This moral dimension of scientific credibility economies was further elaborated on by Daston and Galison (1992, 2007; also see Daston, 1995), especially in the context of iterations of objectivity, as vehicles of credibility, in their historical work on medical imaging. They refer to this as moral economies of science. In the context of proper science this moral dimension is, for instance, indicated by terms like responsible conduct, research integrity or 'good science' and 'good research' more generally (Bartscherer, 2026). The value saturated nature of the discourse about 'good research' and improper conduct, can further be seen in so-called tone debates (Derksen & Field, 2022), or conceptualisations of (current) science reform as a civilising process (Penders, 2022) in the context of language that is perceived as abrasive and acts of shaming (also see other Open Science related conceptualisations of moral economies: Bacevic & Muellerleile, 2018; König et al., 2024). Meskus and co-authors (2018) remind us of the Greek etymology of the term economy, roughly meaning household management. As such credibility work is the researchers' management of their own and their research communities' households in interaction with different

audiences. Daston and Galison (1992) also call the moral economy of mechanical objectivity noninterventionist objectivity, because it is characterised by attempts to protect scientific knowledge production from the supposedly corrupting effect of human intervention and bias to keep procedures and resulting knowledge claims clean (Daston & Galison, 2007; Field & Derksen, 2021; Penders, 2022). Similar to Collin's (1975) description of algorithmic replication models, the moral economy of mechanical (noninterventionist) objectivity seems to go hand-in-hand with procedural interventions aiming to secure and improve research integrity, by providing guidelines and instructions for conduct or aiming to make the (scientific) community virtual witnesses (see e.g., Field & Derksen, 2021; Meskus et al., 2018; Penders, 2022; Steneck, 2007), through for instance more transparent and detailed reporting and preregistration, as well as the process of sharing material.

Application of the framework in the dissertation

In the first two empirical chapters on research integrity and research ethics in the lab I use the framework for understanding whether and how these institutional and regulatory notions of 'good research' are represented within the research groups as well as which role they play for the researchers' management of credibility, with their own perceptions and conditions for 'good research'. As such, the first two chapters explore whether and how the concepts of research integrity and research ethics, move from the left side of figure 1 to the right side and what type of negotiations take place between the notions of ethicality and 'good research' on the right side and those official notions on the left. In the last two empirical chapters on preregistration and replication or reproducibility in the lab, I use the theory to frame the role that these two proposed procedural interventions for the improvement and assurance of research integrity play in their credibility work as well as how they fit into their existing workflows. As such, the last two chapters are mainly situated on the right side of figure 1 as research communities' internal negotiations and enactments of practices that demonstrate forms of 'good research' and are used for the credibility work and management, as part of their specific epistemic cultures (Knorr-Cetina, 1999; Shapin, 1995a).

Chapter 4:
The ethnographic methodology

In this chapter I outline the ethnographic methodology I employed throughout the study. I will first shortly motivate the ethnographic methodology followed by a short introduction of the notion of 'ethnographies of science' to give some required context and situate the study in the larger laboratory studies literature. Next, I will justify why meetings in laboratories are an opportune site to capture notions of 'good research' as well as enactments of research integrity 'from within' research communities as well as groups and laboratories. Afterwards, I will elaborate on the constructivist epistemology underlying this work and go over to my positionality statement. Subsequently, I outline the data collection procedure followed by a short excursion into the process of getting access. Then I will provide details about the field observations and interviews, describe the analysis and end this chapter with some ethical considerations.

Motivation

While there are endeavours to recognise and investigate epistemic diversity in the context of the enactment and negotiation of research integrity (see e.g., Davies & Lindvig, 2021; Hastings et al., 2023; Penders et al., 2009; Valkenburg et al., 2021), recent attempts at capturing researchers' notions of 'good research', research integrity or improper conduct are often based on problematic units of analysis, with focus on high level field distinctions like humanities, natural sciences, social sciences, medical sciences, and technical sciences (see e.g., Sørensen et al., 2021; Ravn & Sørensen, 2021; for exceptions see e.g., Brenninkmeijer & Pols, 2026; Pols et al., 2024; van Drimmelen et al., 2024, 2025). However, as Laudel (2024) put it in the context of similar work on field specific notions of research quality: 'None of these units of analysis corresponds to scientific communities that jointly advance a shared body of knowledge' (p. 4). Hence, they tell us little about what is actually going on and valued in working research communities or groups where knowledge is produced between people that depend on each other for the collective process of scientific knowledge production (Whitley, 1984). Furthermore, a lot of the research on research integrity, whether it is interview, focus group or survey based (Ravn & Sørensen, 2021; Sørensen et al., 2021; Labib et al., 2021), requires researchers to enact or translate notions of 'good' or 'bad research' to scholars outside of their research community (Gläser, 2024). Therefore, those are hardly the researchers' inside notions of 'good' or 'bad research'.

Based on the current literature and for investigative purposes I treat research integrity and research quality as (closely) related concepts (see e.g., Helgesson & Bülow, 2023; also see Chapter 2). Research quality is even part of the first principle of research integrity in the ALLEA, which is about the reliability in ensuring quality of research. The most recent version of the European code of conduct, for instance, states: 'Research integrity [...] encompasses the basic responsibility of the research community to formulate the principles of research, to define the criteria for proper research behaviour, to maximise the quality, reliability, and robustness of research and its results, and to respond adequately to threats to, or violations of, good research practices' (ALLEA, 2023, p. 3). Moreover, it seems naturally contradictory to claim that one's research has high integrity when it is of low quality. Gulbrandsen (2000) characterised research quality as a tacit concept, meaning that researchers' understanding and knowledge of 'good research' is mainly tacit, making especially surveys inappropriate to capture them. I conjecture that the same applies for 'good' or 'bad research' in the context of research integrity where researchers might not be aware of or unable to explicate how they practice or talk about notions of research integrity, because they are just part of the lived experience, tacit or trivial rather than explicit and worth mentioning (for them) (see e.g., Hangel & Schickore, 2017). Therefore, I employ an ethnographic methodology to make the tacit, taken for granted and invisible visible as well as understand research integrity from the researchers' perspective (Hammersley & Atkinson, 2019; Latour & Woolgar, 1979; van Drimmelen et al., 2024). Due to the closeness of the meanings of research integrity and research quality, both conceptually (theoretically) and in practice, I assume that the concept and notions of research integrity can be studied from within as well through a combination of field observations and interviews (Helgesson & Bülow, 2023). Moreover, if the aim is to capture the perceptions and values of the actual researchers, not just for its own sake, but also to be able to investigate how they relate to and agree or disagree with perceptions and values as represented in the research integrity literature and guidelines as well as how recommendations for improved or research with integrity, like data sharing, reproducibility, and preregistration are experienced and practiced in the everyday lives of actual researchers, we need to study the negotiation of research integrity and enactment of 'good research' from the inside of local sites of knowledge production (Brenninkmeijer & Pols, 2026; Davies & Lindvig, 2021; Knorr-Cetina, 1995; Pols et al., 2024; Valkenburg et al., 2021).

Importantly, in contrast to classical ethnographies of science, although these components of knowledge production and their context are intrinsically linked to issues surrounding research integrity, in this study the focus is less on how and what knowledge is produced, but rather on how researchers talk

about and engage with aspects of their work that are or have been related to research integrity in the existing literature (see e.g., Labib et al., 2021; Sørensen et al., 2021). In other words, the aim is to capture expressions and interactions related to research integrity from within the lab. Hence, the focus is more on aspirational values or a commitment-based approach for capturing what research integrity is to the actual researchers that build the basis for a thick ethos of research integrity (Vincent & Hooper, 2025), and later it can be explored how that relates to what is found in the guidelines, courses, and wider research on research integrity literature (Labib et al., 2024, Valkenburg et al., 2021). I do not want to present the researchers with conceptual containers based on preconceived notions of 'good research' and ways to produce as well as justify knowledge from an outside perspective into which they must fit or translate their own inside notions of 'good conduct' and knowledge production or justification (Müller et al., 2024; Pols et al., 2024). For that reason, I do not precisely define or operationalise research integrity or 'good research' and keep it intentionally open for the researchers to fill (Pols et al., 2024). The policy's and research on research integrity's outside perspective of their research and their ideals of proper conduct might not be representative of their actual ideas and ideals, thereby potentially creating a mismatch or imbalance between actual research and researchers' values and research policy as well as research on research integrity alike (Labib et al., 2024; Valkenburg et al., 2021; Vincent & Hooper, 2025). Actually, understanding researchers should come prior to change or even improvement, and in this case even definition and operationalisation. Therefore, we need to investigate the notions bottom-up from the inside to actually be able to understand their epistemic culture and terrain. To that end I employed a combination of field observations during meetings in the labs and semi-structured interviews.

Ethnographies of science

Following and to a certain degree building on the philosophical, theoretical, and anecdotal works of scholars like Ludwig Fleck (1935), Thomas Kuhn (1962), Michael Polanyi (1958), Imre Lakatos (1970), and Paul Feyerabend, (1976), concerning the contingency of science regarding the historical, political, social or generally non-epistemic factors influencing scientific knowledge production, the ethnographies of scientific laboratories started in the 1970s in Science and Technology Studies (STS) (see e.g., Garcia-Sancho & Knorr-Cetina, 2018; Lynch, 2024).[3] The emergence of laboratory studies was part of a

[3] I am doubtful about whether the social and epistemic factors of (scientific) knowledge production can actually be separated (more of a conceptual lens or simplification than practiced reality), but these authors stressed that purely epistemic (cognitive or rational) considera-

development in the social studies of science at the time, attempting to go into the laboratory and open-up the 'black box' of scientific knowledge production (see e.g., Shapin, 1988; Whitley, 1970). Before the emergence of traditions like the sociology of scientific knowledge (SSK) in the 1970s with scholars like Michael Mulkay, Nigel Gilbert, Steven Shapin, Harry Collins, Michael Lynch, Steve Woolgar, Bruno Latour, Richard Whitley, David Bloor, Karin Knorr-Cetina, Steven Yearley, and Barry Barnes the sociology of science was to a large degree comprised of the American Mertonian structural-functionalist approach (Garcia-Sancho & Knorr-Cetina, 2018; Pickering, 1992; Shapin, 1995b, 1988). This literal move into the lab by science studies scholars in the 1970s can be seen as an emic turn in the history of science studies (see e.g., Knorr-Cetina, 1983, 1995; Pickering, 1992; Shapin, 1995b; Whitley, 1970). Very generally speaking the ethnographic studies of science, using (immersive) methodologies from sociology and anthropology, aimed to investigate the social construction of scientific knowledge under the conditions of local sites of practice (Stephens & Lewis, 2017; for cultural approach to science studies see Knorr-Cetina, 1995). Hess (2001) distinguishes between two generations of ethnographies of science in STS, also called laboratory studies (Knorr-Cetina, 1983). Accordingly, the first was mainly focused on the processes and practices underlying the social construction of objective verified scientific knowledge and the second generation was more focused on the role of cultural and political conditions in scientific knowledge production (Hess, 2001)[4]. Classical examples of lab ethnographies are the influential works of Lynch's (1985) *Art and Artifact in Laboratory Science* based on an ethnographic study in a neuroscience laboratory, Latour and Woolgar's (1979) *Laboratory Life* based on a stay in a neuroscience laboratory, Sharon Traweek's (1988) *Beamtimes and Lifetimes,* presenting an ethnographic study of high energy physics, and Knorr-Cetina's (1999) *Epistemic Cultures* based on a comparative lab ethnography between high energy/ particle physics and molecular biology (Stephens & Lewis, 2017)[5]. This tradition of science studies is more or less continued in current ethnographies of science (Lynch, 2024). Some recent examples

tions are insufficient to describe and understand the workings as well as the evolution of science. Which is a sentiment I tend to agree with, you need both the epistemic and the social to explain and understand research and knowledge production more generally (see e.g., Bourdieu, 1975; Whitley, 1983, 1984).

[4] Knorr-Cetina further distinguished the first generation into those laboratory studies investigating controversies concerning knowledge claims and how legitimacy was established (see e.g., Collins, 1975), and those that investigate science in the making (Stephens & Lewis, 2017).

[5] Lynch's ethnomethodological study was actually done before Latour and Woolgar's (1979) Laboratory Life, but the publication of Art and Artefact was significantly delayed (see e.g., Lynch et al., 2019).

are for instance Janet Vertesi's (2020) *Shaping Science*, based on a comparative ethnography between two space missions at NASA, Bechky's (2020) *Blood, Powder, and Residue,* based on the ethnography of a crime laboratory, Peterson's (2016) ethnographic study of the research standards and tacit aspects in infant and toddler research in developmental psychology towards the production of statistical significance, and van Drimmelen's ethnographic study in end-of-life research about researchers' discretion (degrees of freedom) during various research processes and the associated uncertainty surrounding them (van Drimmelen et al., 2024, 2025).

For illustrative purposes I will shortly outline what Latour and Woolgar (1979) did in *Laboratory Life: The Social Construction of Scientific Facts* and what Michael Lynch (1985) did in *Art and Artifact in Laboratory Science: A study of shop work and shop talk in a research laboratory*, in a bit more detail.[6] Latour and Woolgar's (1979) investigation was done with participant observations, interviews as well as document analysis of manuscripts, drafts and exchanges (letters) between researchers.[7] Their material was based on Latour's 21 months stay (October 1975 – August 1977) as a part-time lab technician in Roger Guillemin's neuroscience laboratory at the Salk Institute, giving him close and free access to nearly all practices and artefacts of the laboratory. Similar to typical anthropological ethnographers, Latour tried to immerse himself into the lives of a foreign tribe (see Becher, 1989), he saw himself as a naïve observer (anthropologist) treating the scientific lab like a foreign culture (Latour & Woolgar, 1979). As they put it: 'Scientists in our laboratory constitute a tribe whose daily manipulation and production of objects is in danger of being misunderstood, if accorded the high status with which its outputs are sometimes greeted by the outside world' (Latour & Woolgar, 1986, p.29). They argue that the outsider or 'stranger' perspective allows them to make sense of an 'alien culture' and gain insights into 'aspects of culture taken for granted by its members' (Latour & Woolgar, 1986, p. 278). Hence, it can allow for the identification of tacit decisions and tacit knowledge components in scientific knowledge production and maybe even for the explication of collective and relational tacit Knowledge, according to the authors (Collins, 2010; Latour & Woolgar, 1979; also see van Drimmelen et al., 2024). Due to the fact that Latour conducted participant observations as a part-time technician in

[6] In subsequent editions the subtitle was changed to just: The Construction of Scientific Facts

[7] However, during that time these kinds of anthropological and ethnomethodological approaches did not have to conform to modern bureaucratic and standardized ethical restrictions of today that are largely based on non-ethnographic and non-anthropological notions of proper conduct written down in guidelines (see for instance Herzfeld, 2023; Moors, 2026; Shore, 2024 for how ethnographic research can clash with ethical guidelines and associated notions of research integrity).

the lab and that they attempted to construct a reflexive ethnography of the Laboratory, the report provided in Laboratory Life is also, to a certain degree, based on an autoethnography: 'Our particular use of an anthropological perspective on science also entails a degree of reflexivity not normally evident in many studies of science. By reflexivity we mean to refer to the realisation that observers of scientific activity are engaged in methods which are essentially similar to those of the practitioners which they study' (Latour & Woolgar, 1986, p. 30). In other words, in their account of the construction of scientific facts they draw from a combination or back-and-forth between their own experience of participating in the laboratory life and their observations of it. Resultingly, they conceptualise science as the process of (socially) constructing facts by making order out of chaos (also see Knorr-Cetina, 1983). Part of that chaos is a universal availability of alternative but justifiable interpretations or perceptions of scientific data and the omnipresent fallibility of such perceptions or interpretations (Latour & Woolgar, 1979; for recent examples see De Vries et al., 2006; Gelman & Loken, 2014; van Drimmelen et al., 2024, 2025).

Michael Lynch's (1985) *Art and Artifact in Laboratory Science: A study of shop work and shop talk in a research laboratory,* presents a report based on visits in a neuroscience university research lab. He visited the lab several times a week, over a six-month period. After that for one and a half years Lynch (1985) continued with less frequent visits to the lab. Access to the lab was achieved through existing social contacts (which is often how ethnographers get access to sites, see Hammersley & Atkinson, 2019). The stated goal of Lynch's ethnographic study was to capture the 'analyzable social basis in the local conditions' of the 'social accomplishment of natural scientific order' (Lynch, 1985, p. 1; also see Knorr-Cetina, 1983). Lynch decided to focus on one specific case in the lab — the procedure of axon sprouting — because he thought it is impossible to observe everything that is going on in the lab and such a focus would facilitate a more in-depth comprehension of the specific shop talk and shop work in relation to that specific procedure in the lab (Amann & Knorr-Cetina, 1989; Hammersley & Atkinson, 2019; Lynch, 1982). During the stay he took field notes, kept a journal and recorded some of the talk he was able to observe in the lab. While his note taking did not have a particular structure, he did specifically focus on aspects that were tied to the local conditions and expressions of uncertainty by lab members (Lynch, 1985). Furthermore, he wrote down what struck him as interesting and what was hard to understand for him in what he observed. Lynch (1985) indicates that he did not achieve contributory expertise during his stay, but he was immersed enough to attain (some) interactional expertise enabling him to understand and get access to the technical talk, practices and social interactions of the lab (Collins & Evans, 2002, 2007).

Another motivation behind Lynch's (1985) approach, and many proponents of the sociology of scientific knowledge as well as cultural approaches to studying science, was a dissatisfaction with the previous reliance of science studies on scientific reports, which only provide an outside – etic – perspective of science (Lynch, 1985). He suggests that scientific reports (usually) omit parts of the actual practices of science in the lab and with an ethnography of the lab, similar to Latour and Woolgar (1979), he attempts to uncover potential differences between clean and ordered versions of science in reports (articles and books) and science as actually practiced within the lab under specific local conditions (also see Knorr, 1977; Pickering, 1992). Hence, maybe not going as far as Medawar (1963) and likening the scientific paper to fraud, he thought that the scientific reports in publications do not tell the full story and can obfuscate what really happens in the lab (Lynch, 1985).

While shop work are the practices and activities in the lab, shop talk is the dialogue and talking by researchers, that accompanies their work in the lab as crucial components of knowledge production (Amann & Knorr-Cetina, 1989). This can be critical and constructive discourse with other lab members, but also self-talk where researchers think out loud (Amann & Knorr-Cetina, 1989; Lynch, 1982, 1985; Philipps & Paruschke, 2024). As such shop talk is situated within the local process of knowledge production, prior or in contrast to written reports adapted to formal publication requirements and considerations about how to convince the research community of one's work (see e.g., Lynch, 1982; Schniedermann, 2025; Shapin, 1995a). Next to the talk accompanying the everyday work of researchers, is the crucial role of meetings (Knorr-Cetina, 1981; Latour & Woolgar, 1979; Lynch, 1985; Philipps & Paruschke, 2024; Stephens and Lewis, 2017).

Meetings as an entry point

As for other types of groups and organisations, regular meetings are a crucial site where the co-production and negotiation of the socio-epistemic order of the lab takes place (Jasanoff, 1996; Philipps & Paruschke, 2024; Sandler & Thedvall, 2017; Stephens & Lewis, 2017). Philipps and Paruschke (2024) even go so far as to state that 'ordered talk in regular internal meetings is, like shop talk, constitutive of doing science' (p. 260). Therefore, episodic meetings can provide a window into the local epistemic cultures of research groups or even individual collaborations (Laudel, 2024; Müller et al., 2024; Stephens & Lewis, 2017; Philipps & Paruschke, 2024; Sandler & Thedvall, 2017). In their observations of, mostly online, lab meetings of an optical physics research team, Philipps and Paruschke (2024) describe episodic meetings as a space where researchers give presentations, discuss their work, give feedback, but also for announcements and supervision. Therefore, meetings and seminars

are also a site for undergraduate and graduate students' socialisation into the research community (Delamont & Atkinson, 2001; Müller, 2025; Philipps & Paruschke, 2024). All in all, regular lab meetings are contributing to the management of the focus and direction of the research group, the assurance of the quality of the research as well as the output and generating or inspiring new research (Philipps & Paruschke, 2024).

In the context of a distinction between an outside or etic and an inside or emic (research practice-based, see Laudel, 2024 or Pickering, 1992) view on research quality 'from within' (Müller et al., 2024; Langfeldt et al., 2020; also see Brenninkmeijer & Pols, 2026; Pols et al., 2024, for emic perspectives on research integrity), Müller and colleagues (2024) ask: what might be situations or locations of research activities that could function as entry points to investigate inside notions of research quality in the humanities? They explicitly mention lab meetings and supervision meetings as entry points to studying notions of research quality from within the humanities. Furthermore, recently the role of supervision for research integrity has gained more attention in guidelines and research on research integrity (see e.g., ALLEA, 2023; Haven, 2025; Haven et al., 2023; Mejlgaard et al., 2020). Meetings where researchers discuss their research can also be seen as representative of social interaction for the microanalysis of social order during ethnographies (Müller et al., 2024; also see Müller, 2025 for seminars as 'quality sites' in the humanities). Goffman defines social interactions as: 'that which uniquely transpires in social situations, that is, environments in which two or more individuals are physically in one another's response presence' (Goffman, 1983, p.2). Laudel (2024) investigated bottom-up practice-based field specific notions of research quality in plant biology and medieval history with, among other things, field observations during, informal assessment situations, among them local research seminars (e.g., journal clubs and lab meetings).

Epistemology

Epistemology is the theory of knowledge that builds the foundation for the way of knowing underlying and guiding the methodology of a study, it determines how we produce as well as evaluate knowledge and what is (assumed) knowable (Carter & Little, 2007; Ulpts & Schneider, 2025; Willig, 2013). Like classical laboratory studies (see e.g., Knorr, 1977; Lynch, 1985; Latour & Woolgar, 1979), this research follows a constructivist epistemology rather than a positivist one. Therefore, in contrast to positivist epistemologies, it is not about discovering some underlying truth but about understanding how notions of research integrity or 'good research' as well as related concepts are formed and used by actual researchers in their everyday activities and in accordance with their specific environments (Moon & Blackman, 2014). Furthermore, I do not

take an objectivist stance, that presupposes the ability to discover some underlying often singular truth that exists independent of me or the researchers that I observe. Similarly, Laudel (2024), based on Gläser's (2024) culture-based framing of research quality, in her investigation of field specific notions of research quality distinguishes between a constructivist understanding and the more prevalent essentialist conceptualisations of research quality (also see Guston's (2000) justification for using Gieryn's constructivist concept of boundary work in the context of research integrity). Likewise, here the enactment of research integrity is captured as the result of repeated interactions between researchers within a specific research community (group) and in this case also to a certain degree the outcome of the interaction between the observed researchers and me (see e.g., Gläser, 2024). This epistemological position is somewhat similar to standpoint epistemology, outlined by Donna Haraway (1988) in her description of situated knowledge as being contingent on local conditions of knowledge production (also see Knorr-Cetina, 1983, 1995; Lynch, 1985). She concretises the inherent subjectivity of research and rejects the assumption of an objective view or a view from nowhere. Unlike in positivist research subjectivity is not treated as a vice but is conceived through reflexivity as a virtue that is an inevitable part of the knowledge production process (Braun & Clarke, 2023; Field & Derksen, 2021; Hammersley & Atkinson, 2019; Willig, 2013; Yadav, 2022; for the role of reflexivity for the credibility of ethnography see Aunger, 1995 and Pachirat, 2018). Likewise, as the ethnographer I do not have a view from nowhere but a 'partial view' (Haraway, 1988, p. 590) based on my specific positionality (Pachirat, 2018). The knowledge and findings instead of being discovered are rather co-produced by the participants and myself during the interactions of the fieldwork and the interviews. Therefore, a different observer in different or even the same lab might capture different experiences and notions of research integrity than I. Hence, instead of passive (scientific) voice, this dissertation is mostly written in first person to demonstrate my role in the research as an active component of the knowledge (co)production process that I describe in this manuscript.

As is long acknowledged in most qualitative research approaches, an important aspect affecting the research and knowledge co-production is the kind of relationship that the researcher has with the participants (Hammersley & Atkinson, 2019). This is often known as trust and building rapport. Latour and Woolgar (1986) also identify this as a crucial factor influencing their anthropology of science and write: 'it is clear that the kind of information provided by scientists will have a significant effect in shaping investigators' reports and that the information provided depends, in turn, on the nature of the relationship between scientists and investigator' (pp. 19-20). The establishment of rapport, of course, started with the first email, and was further supported

through me joining and introducing myself during one of their lab meetings prior to the fieldwork. Furthermore, part of that relationship between myself and the researchers were also occasional lunches and social events like visits to bars and restaurants.

Positionality statement

I am pursuing a PhD in Science Studies at Aarhus University in Denmark and am trained in qualitative as well as quantitative methods. Prior to my PhD project I did not have any experience with ethnographic methods, but for my bachelor's and master's degree I did interviews and thematic analysis to investigate researchers' experiences and opinions about peer review and the purpose of statistics in their research respectively. As a psychology and social science student at the University of Groningen in the Netherlands I was very quickly confronted with and got interested in issues surrounding reproducibility, self-correction in science and research integrity more generally as well as the proposed solutions as part of current science reforms, like Open Science and metascience, to address these perceived problems in science. While acknowledging that there are certain issues in science that need to be addressed and that some aspects of current science reform can be useful, I am quite critical about broad stroke pushes for Open Science or research integrity without acknowledging the practicalities and complexities of the diverse research landscape and the contexts in question (Nelson, 2021; Peterson, 2021). Furthermore, I am critical about the communities behind metascience and research integrity, because I see a lack of reflexivity and engagement with already long-established science studies dedicated to the investigation und understanding of science (Bartscherer & Reinhart, 2025; Peterson & Panofsky, 2023). I acknowledge that my critical stance and the experiences I had up to this point will inevitably influence the way I conduct my research, and it might even influence the researchers I am investigating in their behaviour and responses. Usually, I tried to let the researchers guide the tone of the interviews, but in case they voiced critical opinions I engaged with such criticism, inquiring, for example, what exactly they are critical about and why or what certain things mean to them.

Data collection

I opted to select research groups for the fieldwork that conduct research on psychology adjacent topics and are disciplinary-wise located close to psychology, for three main reasons:

1. the current ongoing discussions surrounding a 'replication', 'reproducibility' or 'credibility crisis' as well as large portions of the research integrity and Open Science discourse are mainly focused on and originate in fields of psychology and biomedicine (see e.g., Clancy et al., 2026; Nelson et al., 2021, 2022).
2. Psychology and biomedicine seem to have their own community or even sub-discipline dealing with issues and solutions for research integrity called metascience or meta-research (Ioannidis, 2018; Peterson & Panofsky, 2023). Therefore, I conjecture that in laboratories that are psychology adjacent the likelihood is higher that reform ideas and issues surrounding research integrity are addressed or even discussed more frequently.
3. I also opted for psychology related laboratories, because of my own background to ensure that I can comprehend the conversations and already enter the field with a higher chance of enough interactional expertise to ensure proper inquiry (Collins & Evans, 2002, 2007; Lynch, 1985). The time constraints of the project made proper immersion through prolonged engagement in foreign scientific tribes somewhat unfeasible.

Moreover, I chose to observe groups of a specific size, because the larger the group the more data I might be able to gather due to more meetings where researchers discuss their research with each other. Hence, groups of only three or four people might be problematic for the fieldwork, especially considering the time constraints of the project (PhD within 3 years with one year for fieldwork in 3 European countries). However, the probably most dominant criteria determining the fields and especially which specific labs I observed was who actually gave me access to their labs and their daily routines as researchers. This dependence on the generosity of the researchers (scientists) in conducting lab ethnographies was already noted by Latour and Woolgar (1986): 'The choice of laboratory was determined mainly by the generosity of one of the senior members of the institute …' (p. 39). Due to the fact that this study is embedded within the larger *ETHOS – Exploring Research Integrity Policies and Practices in 'the Houses of Science'* project which has a focus on Europe the selection of labs to observe was also limited to (three) European countries. This selection process can of course result in a combination of selection and response bias which might limit representativeness, generalisability as well as transferability (Lincoln & Guba, 1985; Yadav, 2022). For example, those labs that think they are already more attuned to issues surrounding research integrity might be more willing to grant me access than those that do not. Moreover, my own bias towards certain disciplines due to my own background and the western European focus potentially makes any findings less representative and generalisable beyond these cases. However, since this study is based on

an emic case study approach instead of an etic approach, generalisability is not the main concern (Emic/Etic Distinction, 2008). Generally, researchers conducting a lab ethnography or an ethnographic study in general aim to investigate the social processes of a phenomenon that is widely distributed or typical to be able to make general claims (Hammersley & Atkinson, 2019; Yadav, 2022). Although, ethnographies are case studies and as such any claims about generalisability are bound to crucial assumptions and unfortunately often claims about generalisability are inappropriate and would require further research to be justified. While I describe cases and make no direct claims about generalisability, because I have no empirical data necessary to ground or justify any speculation about the representativeness of the findings, having done a multi-site ethnography might increase the likelihood of describing typical cases in some regards that provide insights for the population of psychology related disciplines (Yadav, 2022). In the next section I will outline in some detail, how access was achieved and granted as well as how rejection of access was experienced.

Getting access and entering the field

Following some rejections from potential research groups, that I would have liked to observe, access to the different labs was granted after first contacting the lab leaders of each lab, explaining my intentions and then introducing myself to the whole group during one of their lab meetings by giving a presentation and provide a space for questions. I sent 37 invitations with two rounds of reminders; one sent after two weeks and one after a month. Out of the 37 invitations sent to group leaders, seven agreed to participate with their lab groups, 10 actively declined to provide me access, 19 did not respond, and for one I decided against field observations because after consultations with the group leader I recognised that there were apparently not enough occasions for observations, because they rarely have meetings. Furthermore, one group in the UK allowed me access to their lab meetings while I was already there and doing fieldwork with another group. The leader (PI) of the research group I was already observing put me in contact with another group at the location. Due to the close relation of the groups (e.g., shared meetings and journals clubs) the smaller amount of interaction with and observation of the additional group, as I did not observe supervision meetings in that group, I treat them here as one cognitive science lab in the UK (similar to the two psychology labs in Mickel et al., 2023).

Reasons provided by contacted researchers for declining to participate in my research varied from asking what such research would do for the scientists themselves, not having the bandwidths and time due to a combination of being already at the limits of their own capacities and the logistics and practical work

in addition to the time it would require to accommodate an external person in their lab, ethical, confidentiality and sensitivity issues related to their research, since there are no unobtrusive observers this might put extra pressure on researchers, especially PhD students with potentially detrimental effects, lack of understanding what such field research would entail, for instance, what I would do all day other than enjoy the city as well as not believing that such research has credibility, the approach and design I propose in the invitation is too intrusive and demands too much from the scientists and asking me to adjust my research design in accordance to their reasons for rejection, and groups or departments just ending or dissolving. One lab leader after consulting their research group told me that some researchers in the lab already had experience with researchers investigating research integrity and were participants of such investigations. However, they did not have a positive experience but seemed to have experienced the research integrity investigator(s) as intrusive and ill-informed about the actual research processes. Furthermore, the outcome of their involvement in the study of those research integrity investigators lead to an impression of being misrepresented, because the findings of the research integrity researchers were seen as being accusatory, inaccurate and indiscreet. Following these rejections, in an attempt to gain access by appearing less demanding to researchers, I dropped a previously intended document analysis component of the study, that I initially planned to incorporate or demonstrate triangulation (see e.g., Yadav, 2022), and be closer to classical laboratory studies (see e.g., Latour & Woolgar, 1979).

Field observations

Goffman (1989) describes fieldwork and participant observation as a form of witnessing, he said: 'You're artificially forcing yourself to be tuned into something that you then pick up as a witness—not as an interviewer, not as a listener, but as a witness to how they react to what gets done to and around them' (Goffman, 1989, p. 126). The fieldwork was done from mid-September 2023 until end of June 2024. The research groups had a size of about 10 to 25 people (group/ lab sizes fluctuate constantly). I conducted field observations in weekly lab meetings and supervision meetings, as well as other meetings between researchers that I could observe, like journal clubs (reading groups). The field observations were based on an abductive approach. Accordingly, field observations were guided by a combination of a deductive and an inductive component. The observations were informed by previous literature and findings from existing research into responsible conduct of research and research integrity and especially by research into which aspects of the research are important for research integrity (see e.g., Labib et al., 2021; Mejlgaard et al., 2020; Sørensen et al., 2021). Hence to a certain degree my preconceived

notions of research integrity have influenced what I paid attention to and noted down. At the same time the field determines what I could observe, and the unpredictability of the fieldwork added to the inductive component. Field notes are a mixture of somewhat detailed descriptions and heuristic jottings (Hammersley & Atkinson, 2019; Iphofen, 2013; Moors, 2026). Although usually field work would be longer than only two and a half months at each site to achieve a certain degree of immersion (Goffman, 1989), the practicality of doing fieldwork as part of a three-year PhD and my choice to study multiple sites necessitates shorter durations for fieldwork at each site. Since my own background is in psychology and social sciences, with a masters in the history and theory of psychology and a research master's in behavioural and social sciences with a specialisation in statistics and psychometrics, to potentially shorten the necessary time for immersion, I chose sites where the technical and natural language are already known to me. This also provided me with a certain degree of interactional expertise (Collins & Evans, 2007), enabling informative exchanges with the researchers and proper understanding on my end of the co-production of the knowledge presented in this dissertation. This, of course, implies a trade-off between being familiar with their epistemic culture to understand them and being an outsider to be able to notice the taken for granted or even tacit components of their everyday work. I chose research sites in Germany, the UK and the Netherlands (Lab meetings are usually in English due to large international representation in the research institutes). The specialisations of the respective research groups were cognitive science, biopsychology, neuroimaging (cognitive neuroscience), and personality psychology. Due to the practical time constrains, I opted to observe more than one research group, when possible, to gather more data in an attempt to reach saturation (Naeem et al., 2024; Saunders et al., 2018).

Interviews

At the end of each field stay (end of November to beginning of December 2023, end of March 2024, and end of June 2024) I conducted semi-structured interviews with the researchers to inquire about the motivations and reasoning behind their practices as presented in their shop talk (Amann & Knorr-Cetina, 1989; Hammersley & Atkinson, 2019; Lynch, 1982). I did most interviews in person and some online during video calls (via Zoom). All interviews were recorded, the in-person interviews were only audio recorded, while the online interviews were video recorded. Hangel and Schickore (2017) remind us, in their own study of researchers' conceptions of good research practices, interviews as a method to analyse scientific practice and scientists are contested and seen as problematic, because they often yield sanitised versions of actual practices, due to, for instance social desirability (also see Tavory, 2020), or the urge to

defend science and present an ideology, instead of a depiction of actual practice (see Mulkay, 1976). Interviews generally provide stories about practices and not a direct representation of the actual practices. Therefore, I conducted interviews about motivations and reasoning behind the practices and themes that I was able to observe and especially hear about during my field observations of meetings of and with the interviewees. Hence, the interviews were informed by actual practices and expressions that researchers showed in their day-to-day work and the conversations they had with each other about their research. Lynch (1985) distinguished, in his ethnomethodological study of science, between 'talking science' (p. 155) and 'talking about science' (p. 147). Talking science is the domain-specific communication between researchers within their epistemic community or even only their labs, while talking about science is the process of communicating the content of science to the outside, like in interviews with journalists and people investigating science (myself). While I provide a combination of researchers talking science (field observations) and talking about science (interviews), the problem of responding according to social desirability or presenting an ideology of science cannot be eliminated completely (Tavory, 2020). When possible, the interviews were conducted in the language spoken and preferred by the participants, to enable the respondents as much as possible to express themselves in their own words (language). However, I do not speak Dutch. Therefore, in the UK and the Netherlands interviews were conducted in English, while most of the interviews in Germany were conducted in German. I conducted 30 interviews with PhD students, postdocs and professors (see table 1).

The interviews lasted between 30 and 124 minutes. The main topics that are related to research integrity in the literature that I could observe in the field and inquire about in the interviews were: The meaning of research integrity, research ethics, replication, reproducibility, preregistration, and research climate. The interview guide for the semi-structured interviews was treated like a living document with questions added or taken off depending on what I was able to observe in the field with the specific researcher in question (see Appendix A for complete example interview guides). Some of the guiding questions that I asked most regularly across participants are:

- Sometimes the issue of reproducibility or replication came up. Can you tell me what those concepts mean to you (what is that) and why it is important for your research?
- Another topic that came up sometimes was preregistration. Why does it come up so frequently and what does it do for your research?
- Did you have a course or something in good scientific practice or responsible research conduct?

Table 1: Characteristics of interviewees

Discipline	Title	Country	Gender
Neuroimaging	Postdoc	Germany	Male
Personality Psychology	PhD	Germany	Male
Neuroimaging (1)	PhD	The Netherlands	Female
Neuroimaging (1)	Principal investigator[a)]/professor	The Netherlands	Male
Personality Psychology	Postdoc	Germany	Female
Personality Psychology	PhD	Germany	Male
Neuroimaging (2)	Postdoc	The Netherlands	Male
Neuroimaging (1)	Postdoc	The Netherlands	Male
Neuroimaging (1)	Postdoc	The Netherlands	Female
Neuroimaging (1)	Postdoc	The Netherlands	Female
Neuroimaging (1)	Postdoc	The Netherlands	Female
Neuroimaging (2)	Principal investigator/professor	The Netherlands	Female
Cognitive Science	Lecturer	The UK	Male
Cognitive Science	Reader	The UK	Male
Cognitive Science	PhD	The UK	Male
Cognitive Science	PhD	The UK	Female
Neuroimaging	PhD	Germany	Male
Neuroimaging	Postdoc	Germany	Female
Neuroimaging (2)	PhD	The Netherlands	Female
Personality Psychology	Postdoc	Germany	Male
Personality Psychology	Principal investigator/professor	Germany	Male
Neuroimaging	Postdoc	Germany	Female
Neuroimaging	Principal investigator/professor	Germany	Female
Biopsychology	Lab Leader	Germany	Male
Cognitive Science	Principal investigator/ Reader	The UK	Male
Neuroimaging (2)	PhD	The Netherlands	Female
Cognitive Science	Postdoc	The UK	Female
Cognitive Science	Postdoc	The UK	Male
Biopsychology	PhD	Germany	Female
Biopsychology	Principal investigator/ Professor	Germany	Male

Note: Principal investigator here means those are the leaders of the respective groups that I observed.

- How do you in the group talk about such things?
- Are there things in the institute or the direct environment that you would change or where you think this is something that makes it harder for me to conduct good research?
- When you read or hear the term research integrity, what comes to mind?

Interviews were transcribed with help from the transcription software Whisper (Transcriber 1.4) and checked by me to ensure proper transcription. As is typical for qualitative research when anonymity and confidentiality are an issue, during the transcription process identifiers like names, locations, and institute or department names were removed (Saunders et al., 2015).

Analysis

The fieldnotes and transcripts of the interviews were analysed using the coding software NVivo (Version 14.24.2). I conducted a mostly inductive thematic analysis on the field notes and the interviews (Braun & Clarke, 2006, 2019; 2023). However, a certain abductive aspect characterises the analysis, since the focus and coding were informed by existing literature on features relevant to research integrity (see e.g., Labib et al., 2021; Mejlgaard et al., 2020; Sørensen et al., 2021) as well as my own positionality and interests. Thus, the purpose of the thematic analysis was to identify topics and themes in the day-to-day practices of researchers as they pertain to research integrity. Already during the fieldwork, the notes were coded to inform the interview, but also to guide the ongoing field observations. Hence the coding of fieldnotes and interviews was an iterative process. Some of the themes that immediately emerged due to them regularly being part of the conversation between researchers are topics like preregistration, replication, reproducibility, Open Science, and research ethics. I coded the interviews through numerous rounds in consultation with the fieldnotes. Throughout the process some themes were split, and some were merged multiple times until a coherent pattern of themes emerged (see Appendix B for the full codebook).

Ethical considerations

The data is stored on a secure drive at Aarhus University. Identifiable information, like names, institutes or locations, in the fieldnotes and transcripts were deleted to protect anonymity and confidentiality. Although, I acknowledge that the sharing of data in the form of transcripts or fieldnotes can, up to a certain degree, be valuable for issues surrounding accountability, transferability, reproducibility, credibility and dependability of the research and

findings (see e.g., Lincoln & Guba, 1985; Murphy et al., 2021; Yadav, 2022), I chose not to share transcripts or fieldnotes publicly or make them available for further use to protect the anonymity of the participants for ethical reasons and keep the data as confidential as possible (Prosser et al., 2024; Moors, 2026). This is also driven by the motivation to respect the participants wishes, because some explicitly asked before I entered their labs and before they provided answers to some of the questions during the interviews, whether this is confidential and anonymous. However, deductive disclosure can probably never be fully ruled out, because people close or internal to the social milieu of the participants will with a higher likelihood be able to identify participants than someone external to that social environment (Tsai et al., 2016). Hence, external confidentiality might be achievable, but internal confidentiality is often more complicated (Tolich, 2004), or even impossible at times. This is further complicated by the necessary compromise between rich descriptions of the research and ethical consideration of privacy and confidentiality (Moors, 2026). Traceability and accountability will be provided as much as possible by thick descriptions, reflexivity and positionality.

Thick description was introduced into the qualitative research vocabulary by Geertz (1973) and over time extended as a practice in qualitative research to describe the authors' interpretations in their local context with the specific emotions and social relationships (see Ponterotto (2015) for a short history of the concept of thick description in qualitative research). To that end, and to somewhat compensate for not sharing material, I provide more and longer quotes and excerpts from the interviews and fieldnotes.

The study was approved by the Institutional Review Board at Aarhus University (approval number: BSS-2023-008/ BSS-2023-009). Participants received a consent form prior to the field work which, if they agreed to participate, was signed prior to, or during the field work as well as a consent form prior to the interviews. Additionally, prior to the fieldwork in the labs, I joined one of their lab meetings for each of the laboratories to introduce myself, explain what my research is about and provide a platform for any questions that the researchers might have. In the consent form as well as during the introduction meetings I especially highlighted that just because they might consent to having me present during meetings does not imply that I must be present during all of the meetings, and that during any occasions where I am observing they can ask me to step out at any time without any explanation required. Hence, in ethnography consent is a rather fluid phenomenon, instead of a clearly delimited event of a binary decision (Hammersley & Atkinson, 2019; Iphofen, 2013; Moors, 2026), as it might be in other research approaches. Furthermore, the degree to which the consent is informed is complicated by the fact that in ethnographic research it is often unrealistic to foresee what the

research will entail for the participants, as the site of research affects the nature of the ethnography and study design often changes over time (Hammersley & Atkinson, 2019; Moors, 2026).

Chapter 5: Research integrity from within: good science in the lab

In this first empirical, chapter I explore the researchers' conceptualisations of research integrity understood as 'good research' or 'science'. 'Good science' can here best be understood as Pols and colleagues (2024) do: 'Our explanation of 'good science' is not, therefore, our judgement as researchers; it comes from our interlocutors' judgements and ideals, as well as the notions of "good science" inscribed in their research, publishing, and teaching practices. These "forms of the good" (Thévenot 2001) may be ethical in nature but also come in different forms and from different registers, such as methodological goodness, good p-values, efficiency, or good working conditions (Boltanski and Thévenot 2006)' (p. 26). I will first outline whether they are familiar with the term or concept and then delineate what they associate with the term by distinguishing between whether they refer to the researcher, the research or the research system (Helgesson & Bülow, 2023).

The concept research integrity with its regulatory origin and status as a policy object within the bureaucratic policy landscape (see e.g., Davies & Lindvig, 2021; Hesselmann & Reinhart, 2024; Roberts et al., 2020), can easily be found in guidelines, recommendations and the research on research integrity literature, but how it relates to actual research practice and how researchers themselves relate it to their realities and everyday experience as scholars largely remains an open question and requires further research (Brenninkmeijer & Pols, 2026; Pols et al., 2024; Valkenburg et al., 2021; van Drimmelen et al., 2024, 2025). Therefore, at the end of each interview, I asked a comprehension question to directly address research sub-question one and inquire what the researchers themselves associate with the term research integrity (German: Forschungsintegrität), what comes to mind for them when they hear or read that term? This question was intended to actively engage them in what Davies (2019) called 'integrity talk' (p. 1240), a specific form of researchers 'talking about science' focused on good or bad research practices. In line with existing research suggesting a certain disconnect between policy and research on research integrity discourse, on the one hand, and researchers' own understandings and perceptions of 'good science' from within actual research communities, on the other (see e.g., Brenninkmeijer & Pols, 2026; Davies, 2019; Davies & Lindvig, 2021; Hangel & Schickore, 2017; Pols et al., 2024; Schickore & Hangel, 2019), the responses from interviewees indicate that the term itself

is not familiar to them. Hence, the concept research integrity is not known, and the term does not seem to be part of their routine vocabulary. This raises the question of, how far the language and content of the bureaucratic and regulatory bodies, in matters of research integrity, is estranged from the day-to-day conduct of research and researchers' own understanding of what 'good science' is (see e.g., Pols et al., 2024).

> It brought up different things. But research integrity is, I don't really know it's a weird phrase. If that is okay to say? (Postdoc Neuroimaging Lab 1, The Netherlands).

> Little comes to mind. I also never read research integrity anywhere. Is that a term that people use? Is that written about?[8] (Postdoc Neuroimaging Lab, Germany).

These two quotes illustrate the general theme across researchers that the concept research integrity is not known. They describe it as foreign and odd, because they usually 'do not come across that term'[9] (Postdoc Personality Psychology, Germany). Thus, little comes to mind because they actually 'never read research integrity anywhere' (Professor Personality Psychology, Germany). This is a crucial issue, because bureaucratic and policy bodies are responsible for issuing the guidelines that define what is to count as conduct worthy of the term research integrity. In the three European countries (Germany, the UK, and the Netherlands) where I did the fieldwork, through a combination of soft law and institutional self-regulation style research governance notions of 'good science' encapsulated in recommendation and guidelines for research integrity and responsible conduct of research either directly or indirectly influence what gets published and funded as well as who gets hired or promoted through a process of top-down regulation and research evaluation (Davies & Lindvig, 2021; Peruginelli & Pölönen, 2024). However, there are also advocacy and attempts to implement harder forms of research governance in the context of promoting and requiring certain notions of research integrity, with attempts at making specific procedures required. (see e.g., Nosek, 2019; Vincent & Hooper, 2025). The quotes also suggest some surprise that they are unfamiliar with this phrase and that the researchers ask themselves why they do not know the concept and even whether they should know it or if they have missed something. This is visible in the questions about whether it is okay not to know the term and if people use as well as write about that concept. Hence, when they are directly asked about research integrity, they seem

[8] Original in German: ‚Wenig. […], wenig vorstellen. Forschungsintegrität hätte ich auch noch nie irgendwo gelesen. Ist das so ein Begriff, den man verwendet? Also, wird da schon drüber geschrieben?'

[9] Original in German: ‚Mir begegnet dieser Begriff nicht häufig.'

startled and at the same time curious about what this thing research integrity might be. Therefore, instead of only relying on whether they have come across the concept, I also once or twice directly asked whether that is a term they themselves are familiar with and use:

> No. Also that is not something I'd consciously think about, instead that is maybe more unconscious. Probably still somehow follow these guidelines, but not consciously. That just doesn't play a big role[10] (PhD Biopsychology, Germany).

Them not knowing official research integrity vocabulary as found in guidelines (also see Davies, 2019; Davies & Lindvig, 2021), is in line with previous research suggesting that research integrity training aimed at conveying the content of such guidelines is ineffective and that research integrity might be more about personal integrity resulting from early life experiences (see e.g., Shaw, 2019; Shaw & Satalkar, 2018). Furthermore, as the preceding quote shows, the process of following the guidelines for research integrity is not described as part of the conscious experience and work of the researchers. Still, it seems to be assumed that guidelines are 'somehow' followed due to a tacit understanding of what it means to have integrity as a researcher and produce research with integrity. Thus, it could be construed as just a form of life that is part of the tacit repertoire that the researchers acquired through socialisation, defining what it means to do 'good science' in action (Collins, 1985, 2010, 2021; Delamont & Atkinson, 2001; Gulbrandsen, 2000; Polanyi, 1958). The latter relates to later comments by a German psychology professor who talks about the role of the spirit of institutions in guiding a scientific community towards research integrity. Relatedly, research integrity is described as 'ill-defined […] in my mind' (PI Neuroimaging Lab 1, The Netherlands) or that it 'encompasses good science' and is 'very broad' (PI Neuroimaging Lab 2, The Netherlands) as well as 'multifaceted'[11] (Professor Personality Psychology, Germany). Therefore, it is not just described as foreign, but also as hard to grasp or explicate, because on the one hand it is vague and strange for them and on the other hand what makes their science good is not simply one thing, but an assortment of different aspects related to and relevant for the research and the specific situation in which the conduct takes place (see e.g., Penders et al., 2009).

Despite their puzzlement and them not knowing the concept behind the term research integrity, when being asked what comes to mind when hearing or reading it, they start to think aloud or begin something like a free associa-

[10] Original in German: ‚Ne. Das ist auch nichts, worüber ich mir jetzt bewusst Gedanken machen würde, sondern das ist vielleicht eher so unbewusst, wahrscheinlich schon irgendwie diesen Richtlinien folgen, aber nicht bewusst. Das spielt jetzt keine große Rolle.'

[11] Original in German: ‚Es ist halt diese ganz vielschichtige Frage, was ist gute Wissenschaft sozusagen, also was sind Aspekte?'

tion process about what research integrity might be. An exemplar of this stream of consciousness is the following quote:

> That you are ready to do robustness analyses, but also that you are willing to share the data, also that you are willing to say and admit that one has made a mistake. […] stand behind my results. That I don't just want to have some articles as long as I am on them. That I am convinced that I only publish if I really belief that's the case. However, I don't actually know what this really means, but those are associations I have[12] (Professor Personality Psychology, Germany).

At first glance it looks like, although the researchers seem to be unfamiliar with the concept research integrity, their attempts to explicate it often resemble values or principles, such as accountability, honesty, and transparency that can be found in guidelines for research integrity (see e.g., ALLEA, 2023, KNAW, 2018; Shaw & Satalkar, 2018). In other words, one might be tempted to say that the researchers' 'integrity talk' and understanding of 'good science' mirrors the guidelines. However, as this chapter and the dissertation as a whole will illustrate the understandings, experiences and perceptions of the researchers are more nuanced, contingent and messy than it might seem (see e.g., Brenninkmeijer & Pols, 2026; Davies, 2019; Davies & Lindvig, 2021; De Vries et al., 2006; Elizondo & Kaltenbrunner, 2024; Hangel & Schickore, 2017; Pols et al., 2024; Schickore & Hangel, 2019; van Drimmelen et al., 2024, 2025).

Therefore, in the following sections I will go through some of the features of research, researchers and the research system that scholars mention and describe as being, important, relevant, related to, or preconditions for research integrity (Helgesson & Bülow, 2023; Labib et al., 2021; Sørensen et al., 2021). Here is a non-extensive list of issues and topics that come up during the interviews that illustrates the plurality and diversity of aspects that researchers associate with research integrity, once the free association process starts:

Perverse incentives, research climate, research conducted as well as possible, research conducted according to current standards, problem of commercialisation, (in)competence, taking data seriously, sample size, replica-

[12] Original in German: ‚Aber ich hätte jetzt auch tatsächlich gedacht gehabt, sowas wie, schon sowas, also wie Open Science auf jeden Fall dabei ist, dass man bereit ist, dass man Robustheitsanalysen durchführt, dass man aber auch bereit ist, die Daten zu teilen, dass man auch bereit ist, zu sagen, dass man einen Fehler gemacht hat, dass man sagen würde, nee, das stimmt jetzt doch nicht so, das sehe ich anders, dass ich sozusagen zu meinen Ergebnissen auch stehe, dass ich nicht nur einfach irgendwelche Artikel haben möchte, Hauptsache ich stehe da drauf, sondern, dass ich, davon überzeugt bin, dass ich nur was veröffentliche, wenn ich mir auch wirklich glaube, das ist so, aber ich weiß jetzt gar nicht genau, was es wirklich bedeutet, aber es sind so Assoziationen, die ich davon habe.'

tion, reproducibility, transparency, Open Science, 'Nachvollziehbarkeit', getting to the truth, follow teaching and guidelines, comprehensibility, owning one's research and taking responsibility for it, hypotheses testing, rigour, honesty, research quality, trust, societal and practical usefulness, critical thinking, data management, data manipulation, fabrication, plagiarism, error detection, ethics (do not harm), robustness, theory, systematicity, taking credit, authorship, preregistration, registered reports, prevent bias, predictive value, no shortcuts, abuse of power, consciously misbehaving (intent), university cooperations/alliances, personality, morality, humility, being self-critical, and conscientiousness.

Most of these terms are about the conduct of research. However, it is important that this is just a list of what is mentioned, as such it does not provide an indication of which is perceived as more relevant for research integrity, the research, the researchers or the research system and institutions. However, some of the statements by researchers below will go into the issue of responsibility when it comes to research integrity. To narrow it down a bit, for space reasons, and because the others can be subsumed under them in the following, I will focus on honesty, taking responsibility, the role of the research system, humility, researchers in a collective, morality and the personality of the researcher, as well as the relationship between Open Science and research integrity. Due to their reoccurrence during my field observations, the central role they play in current discourse about research integrity and related topics, and the insights they offer for understanding researchers' reasoning and motivations behind practices related to research integrity, I will explore issues such as the ethical conduct with participants, preregistration, replication, reproducibility and research climate in separate chapters (see e.g., Helgesson & Bülow, 2023; Kolstoe & Pugh, 2024; MacLeod, 2022; Peels & Bouter, 2023; Resnik & Shamoo, 2017; Shaw, 2019; UKRIO, 2025).

The moral scholar and personal integrity

Somewhat hinting towards the 'bad apples' hypothesis of research integrity and research misconduct, stating that researchers who engage in misconduct are rare and just bad scientists or unethical and dishonest people (see e.g., Lambert & Degn, 2026; Paradeise & Filliatreau, 2021; Vendé et al., 2025), some researchers describe research integrity as being a part or consequence of the researchers' personality (see e.g., Lambert & Degn, 2026; Shaw, 2019; Shaw & Satalkar, 2018; Tijdink et al., 2016). One such notion, provided by the interlocutors, is that research integrity is part of general integrity, as a fundamental characteristic and value-system of a person. Accordingly, the belief is that people with integrity take pride in their work and experience the act of contributing to knowledge as a path to personal fulfilment, because they will

be satisfied and have a clear conscience if they know they have done it the right way. The next scholar goes as far as stating that, as a researcher, doing work with integrity is 'the main component' for satisfaction in one's work. The scholar also explains what this means, by referring to the opposite, and describing how misbehaving, like cutting corners, would take away the satisfaction of being a scientist.

> How you do one thing is how you do everything. I guess research integrity is just a facet of integrity. Integrity is a fundamental part of your life, like self-concept, pride. You can take pride in work that you know is to the best of your ability, you've done with integrity. [...] And if we're aiming for satisfaction from work, then integrity is like the main component of that. [...] If you know deep down that you cut corners or you did something wrong or something doesn't sit right with your conscience, then it's all for nothing (PhD Cognitive Science, The UK).

Here integrity, whether as a researcher or a human being more generally, is inherently linked to having a conscience and being concerned about how one conducts their research. The personality trait of conscientiousness, which entails following social norms, being goal oriented, and being able to control impulses as well as delay gratification (Roberts et al., 2009), has also been linked to self-reported prevalence of questionable research practices (Schneider et al., 2024). Research indicates an inverse correlation between conscientiousness and self-reported questionable research practices (Schneider et al, 2024). This role of the researchers' conscience also comes up when talking about understanding why researchers might misbehave and falsify or fabricate data due to time pressure and personal as well as working conditions (see e.g., De Vries et al., 2006; Elizondo & Kaltenbrunner, 2024). A postdoc talked about how the time pressure that resulted from specific circumstances during the PhD made the scholar understand how and why some people might rationalise outright misconduct:

> I remember one day, I ran some of my first EEG experiments and I knew I was really running out of time. I was sitting in my office, and I ran the analysis, and nothing comes up and I thought, now I understand why people fake data. Now I understand, because I could fake this and no one would know. And I could even tell to myself, okay this is just to get out of this emergency situation and then I'm going to do things right (Postdoc Neuroimaging Lab 1, The Netherlands).

In response, I also asked the scholar why not to misbehave and fake the data? To which the scholar responded:

> I couldn't sleep at night. No, first it just feels, it is wrong. Once in my previous institute I was supervising a master student and the project was fine but [the student] was stressed. And we were struggling to analyse some EEG data and then, I'm sure I know that [the student] said this as a joke, but [the student] said

> oh maybe we could tweak the data. And I look at [the student] in a way that I could see the fear in [the student's] eyes. […] It really feels wrong. Then also I don't know how people can sleep at night knowing that they have fake data out there (Postdoc Neuroimaging Lab 1, The Netherlands).

This also seems to suggest that the role of a researcher's conscience in not committing misconduct is, at least in part, driven by fear of getting caught. Similarly, another scholar says that research integrity is closely related to morality, indicating that the researcher who is acting with integrity and producing research with integrity is a moral scholar (Helgesson & Bülow, 2023). The response allocates such a significant role to the researcher that it also contains the recommendation to support and fund people of integrity (*integere Personen)* to ensure wider research integrity. However, the researcher also complicates the issue by adding the nuance that people of integrity may not be easy to identify, among other things, because integrity is not necessarily uniformly distributed across all areas of life, and in this case, not even across all areas of research a scholar is involved in:

> However, integrity is for me also really somehow more something like morality. […] I have found people of whom I think they do good research, and they mean well with their research, I take it seriously. I think one could also promote (fund) a lot of good research by promoting (funding) people of integrity (integere Personen). However, identifying them is maybe not that easy. And of course, not all people are showing integrity in all things[13] (Postdoc Personality Psychology, Germany).

Another researcher also recognises this close connection between research integrity and the person but also sees the same complication and slight contradiction about integrity not being equally allocated to each area of life:

> Integrity is for me very strongly related to the person. […] And that doesn't just mean not traveling without a ticket on the subway, but also not falsifying your research results. Or not only not falsifying but doing everything to be sure that what you have discovered actually corresponds to what you claim. […] but, of

[13] Original in German: ‚Aber Integrität ist für mich tatsächlich irgendwie eher sowas, etwas Moralisches. Und das deckt sich tatsächlich auch mit meiner Wahrnehmung. Ich lese glaube ich viel Forschung auch so, dass ich Leute gefunden habe, von denen ich glaube, die machen gute Forschung und die meinen es gut mit der Forschung. Ich nehme es ernst. Ich glaube, man könnte auch viel gute Forschung fördern, indem man integere Personen fördert. Aber die zu identifizieren ist vielleicht nicht so einfach. Und natürlich sind auch nicht alle Leute immer in allen Dingen integer. Es gibt auch Leute, deren Arbeit in bestimmten Bereichen finde ich gut. Aber dann gucke ich mal genauer hin an irgendwelche anderen Sachen, die die gemacht haben und bin schockiert so.'

> course, it's very limited. It doesn't matter, if I'm a murderer in my second life, I can still be a good researcher[14] (Professor Biopsychology, Germany).

Taking responsibility and owning the research

Another theme related to the person is that one takes responsibility and ownership for their work. Specifically, research with integrity is understood as the research that one can own up to and take responsibility for, because one is confident of having behaved correctly and not deceived anyone. One could say that this is a form of being willing to be held accountable for one's own research, which is kind of a translation of the principle of accountability and honesty that we find in research integrity guidelines into the actual research environments (ALLEA, 2023).

> The idea of ownership comes up to my head. And maybe that's because I'm thinking of owning your work and taking responsibility for your work. So, the ideas you put forward and the results you put forward, if you are acting with integrity, then you should have no problems taking full responsibility for the implications of those results. [...] And then, if you take responsibility for the work, you put forward, you feel confident about the fact that you're not misleading other people (PhD Cognitive Science, The UK).

In a sense, another researcher goes even further, describing research integrity as being expressed in having no hesitation to stand behind and defend the research one has conducted and reported before anyone, even one's worst enemy, because one believes in it. One might argue that here again the motivation behind research integrity is the anticipation of scrutiny from the research community, whether such confrontation actually occurs or not. This can also be conceptualised as a culture of distrust, which is thought to enable self-correction. The scholar reports on antagonisms experienced between scientists working on the same topics (Bourdieu, 1975). This also complicates and adds some nuance to the role of competition in science because to some degree it contradicts previous research where competition in science is often presented as a barrier to research integrity and one of the drivers for behaviour that

[14] Original in German: ‚weil ich nicht mehr der Jüngste bin, ist für mich Integrität sehr stark mit der Person verbunden. Das ist eben als Persönlichkeit im Finger. Und das beinhaltet da nicht nur, dass man nicht schwarz fährt in der U-Bahn. Sondern auch, dass man ordentlich sozusagen seine Ergebnisse in der Forschung nicht fälscht. Oder nicht nur nicht fälscht, sondern alles tut, um sicher zu sein, dass das, was man jetzt herausgefunden hat, auch tatsächlich dem entspricht, was man behauptet. Und dass man eben, aber das ist natürlich sehr eingeschränkt. Das ist ja völlig egal, ob ich jetzt in meinem Zweitleben ein Mörder bin. Ich kann trotzdem ein guter Forscher sein. Aber das passt, bei mir passt das nicht so richtig sozusagen. Ich bin da noch eher so traditionell getickt (gestrickt).'

deviates from accepted norms, rules, and standards (see e.g., Davies & Lindvig, 2021; Elizondo & Kaltenbrunner, 2024; Pols et al., 2024). The sentiment is that, if the research was conducted with integrity, you can be assured that you are on the safe side when facing the scrutiny of the scientific community.

> Something that you wouldn't have any problems whatsoever showing and defending against anyone. Be it the general public, be a it a specialist, be (it) your arch enemy in research, because we all have […] someone that does exactly the same things that you do in a lab, the other side of the Atlantic Ocean. So, you do have these funny antagonisms and doing good research is doing something that you could show to anyone, and you would be confident to defend against the scrutiny of anyone, because you actually believe in it (PhD Neuroimaging Lab, Germany).

The researcher always acts against the backdrop of the larger scientific ecosystem (Hardwicke et al., 2020), where different structures at various levels intersect (e.g., research group, institute, specific research community, national science system, and global science system). While, with the incorporation of more mundane transgressions of scientific norms and rules, like questionable research practices, instead of a sole focus on misconduct (FFP) (see e.g., Hesselmann & Reinhart, 2024), also occurred somewhat of a shift in the research integrity literature from a 'bad apple' view to a more systemic view with a focus on the role of the characteristics inherent to the science system and associated perverse incentives (see e.g., Downey & Veitch, 2021; Edwards & Roy, 2017; Kennedy et al., 2023; Steneck, 2011; Vendé et al., 2025), the following researcher acknowledges the role of the system, but still sees the main responsibility with the researcher. The interviewee refers to journals and institutes, which admittedly also carry some responsibility, as 'secondary control bodies' that provide 'guiding principles'. However, the actual conduct of research, and the responsibility of ensuring that what is produced is correct, ultimately rests with the researcher(s) (also see Schneider et al., 2021):

> I think you cannot discuss this in such a disconnected way. […] It's a grey area where my own responsibility ends. For example, that's also an area for the journal. Where the journals can, of course, also evaluate the scripts and see if there are any errors in them. […] I would say the core responsibility lies with the researcher […] So, institutes and the other factors you mentioned are more like secondary control bodies. They're like guiding principles, but I think the activity itself comes from the researcher[15] (PhD Personality Psychology, Germany).

[15] Original in German: ‚Also ich finde das ist alles nicht so losgelöst zu diskutieren. Also die Verantwortung als forscher jetzt für mich als forschende Person ist, dass die Sachen die ich geschrieben habe richtig sind dass wenn ich sage, dass ich Fehler korrigiere dass ich nicht Fehler in meine Artikel rein schriebe Skriptfehler berichte also Ergebnisse von falschen

However, the topic of researchers' responsibility inevitably begs the question what makes you responsible and when suboptimal conduct or actual transgression might be the responsibility of the researcher, or what might weaken such responsibility. One critical aspect that comes up as enabling researchers to be responsible for their actions and potential violations of research integrity is competence or actually knowing better. Thus, on the flipside, being blatantly incompetent and not knowing better or, as one researcher put it, 'stupidity' can weaken one's responsibility. In the following quotes researchers identify education and training as the processes that enable researchers to act responsibly, by helping them develop competence and the knowledge to consciously choose to behave well or badly. They thereby also delegate some of the responsibility to the educational institutions such as universities. But it is suggested that once they have properly educated someone and transmitted the necessary knowledge and skills, the responsibility to some degree, shifts from the educational institutions to the researchers (also see Horbach et al., 2024).

> No integrity (nicht integer) would be: I fake/ falsify data, I p-hack; so, I know it only comes out in one of three analyses, and I publish that. That would now not be with integrity (integer). If I actually don't exactly know, I simply publish it because that is what I learned to do, […], then it is maybe just stupidity and not without integrity (nicht integer). Then it might just be bad education. For me I think with integrity that is also a bit about actually being able to decide[16] (Professor Personality Psychology, Germany).

Skripten quasi das sind im Kern so Integrität Sachen. Aber ich finde man kann auch da schon drüber Streiten, das ist so ein grauer Bereich wo hört da meine eigene Verantwortung auf oder zum Beispiel das ist ja auch ein Bereich fürs Journal wo das Journal, die Journals können natürlich auch die Skripte beurteilen und durch schauen sind hier Fehler drinnen. Und ich finde was das störend macht ich weiß nicht, ob das im strengeren Sinne also ja das trägt zu Forschungsintegrität bei, aber ich weiß nicht ob das der Kern davon ist. Das ist eher so eine Kontrollinstanz. Ich würde sagen der Kern die Kernverantwortung hat die forschende Person selbst und das andere sind eher so Kontrollinstanzen. So Institute und die anderen Faktoren, die du gemeint hast, die sind eher so Sekundär-Kontrollinstanzen. Irgendwie so Leitrichtlinien, aber ich finde die Tätigkeit selbst geht ja von der forschenden Person aus.'

[16] Original in German: ‚nicht integer wäre, ich fälsche Daten, ich P-hacke, also ich, ich, ich weiß, es kommt nur in einer von drei Analysen raus, und ich veröffentliche das, ähm, das wäre jetzt nicht integer, wenn ich es aber gar nicht genau weiß, ich veröffentliche das einfach, weil ich das so gelernt habe, ich weiß aber gar nicht, dass was anderes nicht rauskommt, dann ist es vielleicht auch Dummheit, und gar nicht, nicht integer. dann ist es vielleicht auch einfach schlecht ausgebildet. Ja, also ich glaube, bei diesem integer, ist für mich schon so ein bisschen auch eine, überhaupt eine Entscheidungsmöglichkeit zu haben, ist noch so irgendwie, für mich da zumindest drin, kann ich sozusagen, äh, wenn ich es gar nicht besser weiß, ist es dann nicht integer, würde ich glaube ich nicht sagen. Das wäre so ein bisschen der Unterschied. Trotz besseren Wissens mache ich es schlecht. So.'

> Integrity to me really represents not doing obviously morally wrong things like fabricating data. I think that that's, again, it might be that I was lacking with the people that I met on the way, but I haven't really seen issues of that type. But what I have seen more is a lack of training on some things that end up affecting the quality of work and also the quality of life of people. But I wouldn't put that on the side of integrity, because if you think that you're doing things the best way that they can, and even if they suck (Postdoc Neuroimaging Lab 1, The Netherlands).

This further complicates previous conceptualisations of research with integrity being that which you can own up to, because one might also be convinced of the correctness and goodness of one's conduct due to a lack of competence or training. However, if one simply does not know how to do something properly or better, some state, that is not regarded as a lack of integrity. These two quotes bring in the issue of knowingly acting, as a researcher, in a way that would count as not possessing integrity or be bad conduct, which relates to the legal term of intent (Yeo-Teh & Tang, 2024). Therefore, this concerns not only the personal responsibility of the researcher for their work but also connects to the ongoing discussions about criminalising research misconduct and the distinction between honest error, QRPs and misconduct (De Vries et al., 2006; Frank et al., 2026; Resnik & Steward, 2012). It further relates to how issues and regulations surrounding research integrity, proper conduct of research, and misconduct are, or should be, linked to criminological or legal theory and to the legal system in different national and institutional contexts (see e.g., Dal-Ré et al., 2020).

Honesty

Similar to many guidelines for research integrity across the European context (see e.g., ALLEA, 2023; Desmond & Dierickx, 2021), many scholars immediately link research integrity to being honest (also see Shaw & Satalkar, 2018), by defining research integrity as 'how do we do good, open, honest science?' (PI Neuroimaging Lab 2, The Netherlands). However, while the following quote agrees with such a sentiment about the importance of honesty, the scholar does add some caveat to how honest one can and has to be, as specific enactments of honesty in research will vary and depend on the context (Zuckerman, 1984). Although, the minimum remains not to intentionally deceive or mislead. This suggests that the enactment of honesty is situated on a spectrum that locates the conduct of research between the poles of absolute honesty and dishonesty. Which begs the question, what makes a researcher justifiably or not tilde towards either pole? Furthermore, the enactment of honesty is often linked to the practice of reporting research. This becomes especially salient in relation to demonstrating the temporal order of the research, clarifying which

parts are hypothesis-based, and which are exploratory. The perceived criticality of the temporal order of events during the research, and which parts of the research are exploratory or confirmatory, and how this aspect of communicating honesty is enacted in practice, will be explored further in Chapter 7 on the practice of preregistration.

> That you at least not necessarily report everything you have done, but that you try to be as honest as possible. That you at least don't say things that you didn't do, so that you don't lie in your papers. […] at least that it becomes clear what all was tried out with the data and in what way it was planned beforehand, whether it was hypothesis-based or still exploratory[17] (Postdoc Neuroimaging Lab, Germany).

In the following response, honesty is linked to how to treat the data. Here, a researcher suggest that honesty arises from taking the data seriously, indicating something like a responsibility to go where the data leads you:

> Taking my data seriously and being honest about my data (Postdoc Neuroimaging Lab 1, The Netherlands).

As the first quote in this subsection on honesty already indicated, sometimes honesty and transparency are put together or convoluted in line with an Open Science ideology, and linked to issues of reproducibility (Bouter, 2024; Haven et al., 2022). The relationship between Open Science and research integrity will be further discussed at the end of this chapter:

> Doing research in an honest, transparent way that can be reproducible and such (Postdoc Neuroimaging Lab 1, The Netherlands).

Although honesty is often mentioned in the context of data and reporting of the research, honesty is also cited as an aspect of the relationship with and treatment of the participants and the ethical issue of deceiving them, which can be in conflict with internal validity issues like demands characteristics (Orne, 1962), as the following researcher puts it:

> That one definitely, in any case, is also transparent towards the test subjects. And in case they have questions, that they are being answered, then one says okay I cannot tell you beforehand, then at least afterwards at the debriefing talk and say this is what it is all about (PhD Biopsychology, Germany).

[17] Original in German: ‚Naja dass man zumindest nicht unbedingt alles reported was man gemacht hat aber dass man versuch so ehrlich zu sein wie möglich. Also dass man zumindest nicht Sachen erzählt die man nicht gemacht hat also dass man nicht in seinen Papern lügt das wäre gut glaube ich und zumindest das irgendwo halbwegs glaube ich klar wird was man alles ausprobiert hat mit den Daten und in welcher weise das quasi vorher schon geplant war ob das Hypothesen basiert war oder jetzt noch exploratorisch ist.'

Hence, while at face value honesty seems to be one of those values which are treated like truisms and everyone can easily agree on it being a crucial part of research integrity, as the responses indicate, how it is perceived and implemented into practice is actually more subtle, and it can take many forms as well as degrees, according to context and purpose (Moors, 2026). Furthermore, the question of how to enact honesty in specific situations can add uncertainty and carry certain types of moral ambiguity (Moors, 2026; van Drimmelen et al., 2025).

Humility and critical reflexivity

Researchers' presumed humility is an old but not much discussed topic within the sociology of science, or science studies at large. The value of humility is closely related to honesty. Robert K. Merton (1957) already wrote about the scholar's dilemma between the institutional norms of humility and originality (which will come up in Chapter 8 on replication). According to Merton (1957), humility can take on various forms in science, and he highlights two of them: 1) acknowledging the intellectual indebtedness to those who came before you, and 2) acknowledging your personal and scientific limitations. In Merton's own words, there is an 'inner conflict among men of science who have internalized both of them and generates a distinct ambivalence toward the claiming of priorities' (Merton, 1957, p. 647). And as Weingart (2017) put it in reference to Merton (1973): 'the individual scientist finds him- or herself in an ambivalent situation between modesty in presenting truth claims and resoluteness in fighting to prioritize the importance of his or her work. This ambivalence stems from conflicting institutional norms: the main criteria for reward in science are novelty and originality. This is the reason for the emphasis on the certification of priority of discovery. The counter norm is modesty or humility: it stresses the precedence of the universality of knowledge over personal engagement and interest' (Weingart, 2017, p. 112). Intellectual humility or modesty has recently also become a topic for Open Science reformers in their endeavours to enhance the integrity and credibility of research (see e.g., Hoekstra & Vazire, 2021; Ramsey, 2021). However, similar to their lack of critical engagement with earlier sociology of scientific knowledge (SSK) work on replication, current metascientists, research integrity scholars, and Open Science reformers do not critically engage with these older discussions on humility in science (Peterson & Panofsky, 2023). Recently, the expression of humility in scientific writing is also described as the opposite of intellectual arrogance or hype (see e.g., Hoekstra & Vazire, 2021; Weingart, 2017). As a German postdoc states when thinking about what research integrity might mean:

> Investing time in error detection and humility, admitting uncertainty, things like that. [...] the opposite of hype and spin and selling[18] (Postdoc Personality Psychology, Germany).

The researcher also goes on to justify the need for more humility in research and especially in reporting. The reason for the need for such humility, stated, is the inherent uncertainty of research (Elizondo & Kaltenbrunner, 2024; Peterson & Panofsky, 2021; Schickore & Hangel, 2019; van Drimmelen et al., 2025; Whitley, 1984). From own experience, the researcher explains, that one cannot be confident or confident enough to not show humility and be explicit about the uncertainties of the research at hand. Hence, the degree of humility shown by researchers seems to be part of the credibility work that aids in the management of how research is received within the community, or at least by this scholar (Latour & Woolgar, 1979; Geiger, 2021; Shapin 1995a). Therefore, being too confident as a researcher is outlined as an indicator that something is off and a reason to be suspicious and have lower trust in the researcher as well as the research.

> People who are too sure (confident), there I just already have the feeling, they actually have not captured the mindset of a scientist. [...] People who are too early saying that they understood something, that is for me simply antithetical to what a scientist should be. If I have the impression that people embody that and express themselves like that, then that is directly a downgrade for me, regarding how much I trust them. [...] I just know how uncertain a lot of our research findings are[19] (Postdoc Personality Psychology, Germany).

---

[18] Original in German: ‚Aber ansonsten denke ich halt vor allen Dingen an Transparenz, Reproduzierbarkeit, Zeit investieren in Fehlersuche und Bescheidenheit, Unsicherheit zugeben, solche Sachen. Also nicht, das Gegenteil von Hype und Spin und Verkaufen.‘

[19] Original in German: ‚Also für mich ist es ein super Signal, also Leute, die sich zu sicher sind, da habe ich einfach schon das Gefühl, die haben eigentlich das Mindset des Wissenschaftlers nicht erfasst, weil also sobald du dir sicher bist, heißt das, dass du eigentlich quasi ein Prior von 100%, der von Daten nicht bewegt werden kann, du musst eigentlich immer den Hintertürchen offen lassen, nochmal deine Meinung zu ändern für, also, ich meine, also ich glaube jetzt auch nicht, dass Homöopathie sich nochmal bewahrheitet oder so, aber aber das, Leute, die zu früh irgendwie dahin gehen, dass sie denken, so ich ich bin, ich habe es jetzt verstanden, das ist einfach, das ist für mich antithetisch zu dem, was Wissenschaftler sein sollten. Und wenn ich das Gefühl habe, Leute verkörpern das und drücken sich so aus, dann ist es für mich direkt ein Downgrade, wie sehr ich denen traue, so. Ist für mich ein Zeichen von schlechter Kalibrierung, weil ich einfach weiß, wie unsicher viele unserer Forschungsbefunde sind, aus eigener Erfahrung sind Leute, die sich mittelsicher sind, ich denke eigentlich nicht, dass man sich mittelsicher sein kann, vielleicht wenn man so Wahrnehmungsforschung mitmacht oder so, dann kann man vielleicht wirklich teilweise sich sicher sein, aber in dem, was ich so mache, wer sich da sicher ist, ist, liegt falsch. Da bin ich mir sicher.‘

After I followed up about the aspect of being flexible enough to change one's opinion and reflect on the state of the knowledge and assumptions one is making prior, during and after conducting the research, the scholar clarified about the importance of being self-critical or reflexive for research integrity (see. e.g., Field et al., 2024; Valkenburg et al., 2021; Wäscher et al., 2020):

> With integrity it is also the right to be self-critical, I would definitely see that as an important precondition. […] If you think about your own thought you have to be able to also discard things[20] (Postdoc Personality Psychology, Germany).

Hence, part of humility seems to be a self-critical stance as a researcher. Similarly, another researcher also talks about these issues by referring to 'the worst kind of illness' (PhD Neuroimaging Lab, Germany) a researcher can have. This illness is described as a lack of sincerity and critical reflection regarding the research that someone is publishing and representing. The interviewee suggests that some researchers use every trick in the book, without committing actual research misconduct, to defend and support, you could even say, sell their research findings, often without actually being convinced of the validity of those findings themselves. In contrast, the scholar claims, researchers have to be sincere and continuously question the type and degree of validity of the findings they represent. This questioning of themselves is fuelled by being self-critical. Here to a certain degree reflexivity is also visible, since the interlocutor provides a self-diagnosis.

> The worst kind of illness that I have seen and have diagnosed in myself and in other people is sometimes doing something with not being sure whether or not you're finding out it is valid or not. Sometimes you see people that have results that they don't entirely believe in. […] They have plans on how they're going to justify them, and they are not out to trick anyone. […] they would never falsify truth. […] But they employ every other trick in the book to defend something against reviewer number two, and they are miserable because they can come up with nice arguments, but they are not sure themselves. And then I think some people react to that by desensitising themselves. It's almost like they don't have a personal opinion on the truthfulness of their results. […] And this I think comes from desensitising yourself, stopping yourself asking the question […] but I think you should keep on asking it yourself (PhD Neuroimaging Lab, Germany).

The researcher notes that this problematic tendency to blindly defend and promote one's findings appears to be linked to a process in which scholars become detached from, or desensitised to, their own research. Which is clearly

[20] Original in German: ‚Ja. Und ich meine, bei Integrität ist auch das Recht, selbstkritisch zu sein, würde ich da auf jeden Fall auch als wichtige Voraussetzung sehen. Irgendwie nur, dass man ja, wenn man bei jedem seine eigenen Gedanken denkt, der ist Also dann, also man muss ja auch Sachen verwerfen können.'

in line with the previously described component of research integrity entailing taking ownership and responsibility for one's research and finding self-fulfilment and pride in doing research well, because it is so closely linked to one's intrinsic motivation, values and sense of personal integrity (Shaw & Satalkar, 2018). This problematic desensitising oneself from one's own research and not being personally invested in the truthfulness of one's findings, is also being linked to some researchers' orientation towards success in their career and focus on having something publishable. Here, the scholar also hints towards a dilemma researchers have to deal with between doing well career-wise and continually asking oneself whether the research one is conducting is good or not, which turns up again in cases where researchers demonstrate something like a risk-benefit analysis linked to the employment of practices for their management of credibility (Shapin, 1995a). The interlocutor also suggests something like a 'miserable' experience of cognitive dissonance (Festinger, 1957), by striving for validity but simultaneously, as the next quote suggests, also following the contrasting aims provided by a 'mindset of is it publishable or not'. This cognitive dissonance (ambivalence) indicates that there seems to be no complete desensitisation or detachment happening (Merton, 1957).

> You should not get forced into this mindset of, is it publishable or not? You should keep on with the mindset of, is this good or not? Is this true or not? [...] A lot of success, people that are doing well in sciences shift to, is this publishable or not? And they just stop asking the core question. [...] Whereas I think sometimes people that maybe have a harder time separating themselves from the core question... I think these are people that tend to drop out (PhD Neuroimaging Lab, Germany).

The last part of the statement further shows that the researcher seems to see the situation in the science system as so bad that those scientists that do not or are unable to desensitise themselves from the epistemic questions leave academia. Suggesting that the scholar thinks to have a successful career in science, a certain degree of detachment from your science is necessary. Another scholar also talks about the importance of humility for research integrity, in the form of keeping oneself in check, especially concerning hyping, when reporting the research (Hoekstra & Vazire, 2021; Weingart, 2017). However, as with many of the responses from the early career researchers, this might also just be the somewhat naïve and still idealistic outlook of a young PhD student. (This aspect of limiting oneself and being kept in check will also turn up again in Chapter 7 on preregistration):

> Do I really think this shows that? Maybe not, I should tone that down a bit (PhD Cognitive Science, The UK).

However, this self-critical attitude can apparently also become quite extreme where people can even get lost in questioning themselves and their choices, because they are too self-critical and start ruminating about their responsibility while being unable to move forward with their research and they even 'paralyse' themselves (also see Davies, 2019):

> It is important to constantly question your choices. And it's something that if you do too much, it can paralyse you sometimes (PhD Neuroimaging Lab, Germany).

This getting stuck due to self-doubt can be problematic for researchers because in the reality of conducting research in the marketplace of academia, there is always only a limited amount of time. Related to taking being conscientious and self-critical too far in the face of time pressure are a postdocs comments about being somewhat paranoid due to past research experiences and how there is a need to build in controls and limit yourself as a researcher:

> I think I do quite a lot of controls in the data analysis side. But I'm sometimes a bit paranoid. There's a coding error, and then I will wake up, and then I will see that everything is wrong (Postdoc Neuroimaging Lab 1, The Netherlands).

Similar to others above, another researcher relates being self-critical and reflecting on the risks and the validity of the research, as a component of research integrity, to being conscientious about their work. This can, for instance, be demonstrated by staying on top of the game when it comes to the available state-of-the-art techniques and methods and using the best possible approaches or practices (see e.g., Wäscher et al., 2020):

> To be critical like conscientious, just be conscientious about how you do your research. Always think is this is the best way to do this, so always try to stay up to date with analysis techniques or new insights that come in, so for me technically it is trying to do the best. We always thrive to do the most state of the art, things that we know are sound research (Postdoc Neuroimaging Lab 2, The Netherlands).

The role of the research system

In line with recent views on research integrity and potential drivers of transgressions from accepted norms, rules, or standards being systemic, and especially due to the incentive and reward structure of science (see e.g., DeVries et al., 2006; Elizondo & Kaltenbrunner, 2024; Lambert & Degn, 2026; Robinson-Garcia et al., 2023), when talking about the responsibility of the research system for research integrity more directly, scholars state that as previously already suggested, the make-up of the research system might actually be working against research integrity, or even towards the 'reverse' of it. Hence, it is sometimes referred to as perverse incentives (see e.g., Edwards & Roy, 2017;

Steneck, 2011; Stephan, 2012). The next researcher especially points to the ways to be successful in science by being 'liberal' with the data. Furthermore, the researcher indicates that this perception is common in the community and that a change is happening, but at the same time, there is still a lot to do to align the research system with an aspiration towards research integrity:

> We're kind of coming to realise how important this is, because we would like to be in a system where your success as a scientist is correlated (with) your research integrity and now it seems to be the reverse. It's, easy to get ahead just by making up your data or being very liberal in the way you analyse it. I'm kind of happy that those are things, that this is something that has resonated up to awareness in the community relative to how it was before. But we probably still have a long way there (Postdoc Cognitive Science, The UK).

Also referring to academia as a system where what enhances your career and what enhances research integrity might not always be aligned, here a postdoc is referring to the circumstance that as a researcher in the system one has to be at least somewhat strategic in the practices one is using and the choices one is making, because it might have consequences for the career trajectory one can have as a researcher. Hence, we can see something like an investment strategy or cost-benefit analysis concerning the credibility work that scholars engage in (Geiger, 2021; Latour & Woolgar, 1979; Shapin, 1995a). It seems to be perceived, at least to a certain degree, as an either-or situation between research integrity and career. (Davies & Lindvig, 2021; Hangel & Schickore, 2017; Pols et al., 2024).

> I think one also has to be careful, so you don't get so far. I haven't come that far myself yet, but one does not stay in academia if you don't think strategically somehow. If you don't think at all at some point, this project is better for my career. Thus, if you completely ignore such things and are only guided by one's own interest and what is scientifically good, then I don't know how far one gets[21] (Postdoc Personality Psychology, Germany).

More pessimistically, one could be tempted to say that the experience seems to be that to be successful in the current science system, one has to make compromises or even corrupt their values or aspirations and employ practices contradicting research integrity (see e.g., De Vries et al., 2006; Pols et al., 2024). Providing a global perspective on the problematic influence of economic

[21] Original in German: ‚ja, aber ich glaube, man muss auch aufpassen, also ich, da auch, also,, man kommt jetzt nicht so, ich bin jetzt noch nicht so weit gekommen, aber man bleibt jetzt nicht so lange in der Forschung, wenn man gar nicht strategisch denkt irgendwie, wenn man gar nicht mal irgendwie denkt, so, ich, dieses Projekt ist für meine Karriere besser, also wenn man solche Sachen komplett ignoriert, dann, und sich nur von seinem eigenen Interesse und was wissenschaftlich Gutes leiten lässt, dann weiß ich nicht, wie weit man kommt.'

forces in and on science, in the following quote, a researcher describes being somewhat disillusioned about the state of research integrity across disciplines and the globe, due to exposure to a lot of publications in the role of an editor. Here, the scholar indicates that research integrity not just varies across disciplines but also across countries. While acknowledging that this might be a prejudice, the researcher states to be supporter of mostly western traditional scientific research, because other regions have a higher degree of commercialisation of science, which is seen as working against research integrity. One might argue that the scholar is hinting towards processes of epistemic corruption, in which commercial interests exercise influence on scientific knowledge production and dissemination in conflict with research integrity (Sismondo, 2021). More precisely, the scholar likens the research systems in some countries to an environment that treats research like an assembly line, which endangers research integrity, because that 'is not wanted':

> From my perspective there are whole fields [...] and countries where there generally is no research integrity, because that is not wanted. Or just not, so research is seen like producing cars. And, in my view, that is always an environment where research integrity is in danger. I am a supporter of western traditional scientific research, and I think that this might be a terrible prejudice right now, commercialise that too much, also western countries and universities. I think these contexts are actually country specific, in which you can recognise that[22] (Lab Leader Biopsychology, Germany).

While coming up with preconditions for research integrity, the PI of a Dutch neuroimaging lab discusses the issue of the publish or perish culture in science and the nature of publication as a numbers game (Moskovkin, 2024): 'Publishing as fast as possible and more publications is better' (PI Neuroimaging Lab 2, The Netherlands). This means that if you do not publish enough and in specific journals with a certain impact, then you will just be obliviated as a

[22] Original in German: ‚Ich bin Editor in einer Zeitschrift und, sehe deshalb relativ viele Publikationen, sagen wir, schwieriger Qualität und, äh, ähm, ... sowohl vom, also, da können, aus meiner Sicht sind da in ganzen Feldern mitunter auch in, eben, äh, man mag mir da jetzt irgendwelche, das mag ich nicht mit Namen zitiert haben, aber anonym darfst du es gerne zitieren. Das es Felder und Länder gibt, wo Forschungsintegrität generell nicht gegeben ist. Und, äh, weil einfach das als gar nicht erwünscht ist. Oder gar nicht, also einfach Forschung als so wie Forschung gesehen wird, wie Autos zu produzieren. Und, äh, das ist immer ein Umfeld, wo Forschungsintegrität aus meiner Sicht gefährdet ist. Also ich bin tatsächlich so ein Anhänger von, von westlich traditioneller wissenschaftlicher Forschung und ich denke, dass da, ähm, das mag jetzt irgendwie ein furchtbares Vorurteil sein, aber, äh, gewisse Länder, äh ... das zu sehr kommerzialisieren. Auch westliche Länder. Und Universitäten. Das sind tatsächlich, also ich glaube, diese Kontexte sind tatsächlich länderspezifisch. Also. In denen man das aufzeigen kann.'

scientist in the system (this might be a bit different if you have tenure), sometimes also referred to as 'impact or perish' in a metricised market of attention (Biagioli & Lippman, 2020, p. 1).

The PI further describes the role as an editor within such a system and how that relates to a responsibility of ensuring the integrity of the work that is published in the specific journal, while again narrowing in on the issue of a publish or perish culture:

> Publish or perish thing. When I was an editor for [...], we spent a lot of time in writing editorials on what do we think current standards are for stats, animal science, human science, it's different. [...] So, as a journal editor, your task is to think about research integrity. How do you ensure that the quality of the research published in that journal is high? And I just accepted to be an editor for [...] and [...] is spending a lot of time, of course, a lot of energy on the whole publication process and also on the publish or perish disaster (PI Neuroimaging Lab 2, The Netherlands).

Hence, also in the role as an editor these issues like the question of whether replications are needed, and those surrounding statistical methodology or the differences in standards between human and animal research, that are currently part of a reform agenda to improve the credibility as well as integrity of science, seem to be significant enough to be directly addressed in editorials, which can be seen as documents representing the ethos of the journal. While the journal, as a central node in the science system (maybe not in computational sciences), is also described as a body that can potentially and should counter the 'publish or perish disaster' by changing its peer review policy (some for instance recommend a move from a review-curate-publish model to a publish-review-curate model of academic publishing, see e.g. e et al., 2024).

The researcher in a group of researchers

The aforementioned outside perspective towards the research community one depends on regarding what counts as good enough research, indicates that another important aspect of 'good science' or research integrity, next to the role of the individual and the institutions or the system at large, is also the relations and interdependencies between researchers (Whitley, 1984). Hence, the researchers remind us that research is not conducted in isolation, but it is a fundamentally social activity taking place in relation, association and interdependence with other researchers within a collective (as well as other actors, Knorr-Cetina, 1982, 1983).

> I was thinking about a group, for research integrity. How you in coordination with others conduct good research within a system and a group. [...] Or how are

> you a good human and researcher in a group[23] (Postdoc Neuroimaging Lab, Germany).

This is further represented in the following quote. Here, the scholar talks about being able to trust in one's research and represent it to oneself and others, because there is an interdependence between researchers (mutual dependence, see Whitley, 1984). It is mentioned that scientists need to be able to trust or have confidence in the fact that they can rely on the research of others. The mutual dependence and reciprocity of trust is expressed in the lines of: as long as I, as a researcher, can depend on the work and results of others, then they should also be able to trust mine. However, while Whitley (1984) sees fields with high mutual dependence as characterised by standardised control mechanisms and sanctions for transgressions, here it seems to be assumed that a reciprocity of trust ensures the stability and integrity of the knowledge production process of science. This is, again, related to the aspect of taking responsibility for and owning up to one's research. The scholar states that if you act with integrity, you assume others will too. This, of course, might also originate from an assumption that the community specific social mechanisms of control and sanctioning (self-correction), that Whitley (1984) sees as characteristic of high mutual dependence, are working. However, the scholar also adds the complication that there are also critical differences in what people consider research with integrity looks like, or what it means to be a researcher with integrity:

> This is why I think it's all a very trust-based system, which is good. If you trust in your work, you can trust that other people are doing their work as well, and It's taken to their own form of integrity. But I acknowledge that maybe people's notion of integrity (differ), not drastically I think, but like enough to make people's work slightly different (PhD Cognitive Science, The UK).

The following response from a researcher already suggests that the research system is a multifaceted phenomenon which influences the researcher on multiple levels, from the lab group to the institute and the whole global science system. Here, the scholar talks about the community and the system's responsibility to enable exchange and collaboration between researchers with the integration of differing perspectives, but at the same time also to keep potential misbehaviour or errors, like cutting corners, in check, since those could

[23] Original in German: ‚ich stelle mir da sehr viele verschiedene Dinge vor, die sein könnten. Weil eigentlich habe ich zuerst an eine Gruppe gedacht, zur Forschungsintegrität, wie du quasi innerhalb eines Systems einer Gruppe gute Forschung in Abstimmung mit anderen machst. Aber einfach wahrscheinlich aus dem, ich weiß es nicht, Faktor heraus. So klingt es irgendwie, finde ich. Oder wie bist du ein guter Mensch und Forscher in einer Gruppe, so in die Richtung.'

actually harm the community. Indicating that there are limits to the trust in this trust-based system, or that the assumed control by the community is the reason for trusting other researchers in the community. Hence, this might be a subtle reference to the assumed, currently widely doubted, self-correction mechanism in science as the organised social mechanism justifying the trust put in the output of other scientists (see e.g., Ioannidis, 2012; Stroebe et al., 2012), like organised scepticism (Merton, 1942). Thus, we do see comments hinting towards issues concerning the research climate and its importance to the mutual dependence of researchers in their endeavours to produce knowledge (Latour & Woolgar, 1979; Whitley, 1984).

> It needs to be a good community, both within your lab, within your institute, but also on a global scale. I think there need to be the platforms and opportunities to share and exchange with researchers in your field that allow for good science. Not just sit at your own computer and do your own analysis but actually integrate different views and be mindful of the effects that whatever you're putting out, has on other people, both positive when you're sharing and helping other people, but also when you're cutting corners and putting scientific integrity at stake. What down trickle will that effect have on the community? (PhD Neuroimaging Lab 2, The Netherlands).

The next researcher describes their group as being quite into 'good research' and research integrity-related issues, and that talking about mistakes is welcome. This combination of being into matters of 'good research' and having a climate where it is okay to talk about making mistakes is described as leading to a situation where the burden of responsibility for the study, or if something might go wrong, is being shared within the group and not on the shoulders of just one researcher (see Frank et al., 2026 for error culture). Therefore, the researcher states that a good relationship between colleagues and a shared investment in research integrity or 'good research' enables them to share the burden of responsibility for the work as a researcher in a specific lab or group. Moreover, such openness to discuss errors and uncertainties is in accordance with the previously stated theme of being reflexive, humble, and critical about one's research conduct, as a component of research integrity, since it is presented as a kind of precondition or facilitator of reflexivity. This topic of the lab culture will be further explored in Chapter 6 on the ethics in the lab:

> I think we're quite lucky in our group [...], we're all quite into it. We talk about these things. We mention them to each other. It is okay to make mistakes. I think I realised once that I had continued working on a study when the ethics approval had run out or like [PI's] ethics approval, what we thought extended to cover this one didn't. [...] But it doesn't matter. It's ultimately, it doesn't matter. It's not like a personal moral failing. It's just kind of like, whoops, glad we caught that. Let's fix it. It's not like you're seriously blamed for, as if you personally have

> messed up. It's like it was everyone's responsibility. It makes it not that big a deal to mention things to each other because […] it's not like, if you make a mistake, you're a bad person or something (PhD Cognitive Science, The UK).

A German psychology professor links the different levels by talking about the complex interrelations and dependencies between the institutes, the collective and the individual researcher(s) when it comes to the responsibility for research integrity. The professor mentions the necessity for a body above and beyond the researchers that, through an institutional 'philosophy or religion', ensures research integrity. Here, the researcher seems to refer to an implicit or tacit ethos of good research (Vincent & Hooper, 2025). Furthermore, this 'spirit' is described as something that is not reducible to the person. Continuing to state that guidelines are an attempt to capture the aspirations contained in such philosophies (something like a culture of integrity) (Valkenburg et al., 2021). This overarching and tacit 'philosophy' or 'religion' of research integrity might be linked to a PhD's earlier comments about not consciously thinking about research integrity-related things, while they 'probably still somehow follow these guidelines'. Therefore, according to the professor, the individual researcher is the wrong or insufficient stakeholder to focus on; it needs a perspective on the collective and the whole:

> Research integrity is more or at least as much the task of the institute, of the collective of subjects, but also of the institutes, the collective of institutes that are involved. […] One always needs at least the philosophy or religion that keeps up. It plainly doesn't do that with regards to the integrity of research, but what is needed there is something else, super-personal. Guidelines are a bit an expression of what one is actually aspiring, with upright, reliable research with integrity and so on. […] This is actually the wrong measure (benchmark/ standard) in a way because there is only individualised responsibility of the single person. Therefore, it plainly needs a perspective and initiative across personal and collective, maybe considering and developing the whole (Professor Biopsychology, Germany).

The professor goes on and even complexifies the earlier remarks and further elaborates on the relations when it comes to the responsibility of the researcher, the collective and the institute. The professor states that it is problematic to focus too much on only the individuals or only the departments, it requires both. There is even a reference to the role of the aforementioned philosophy or religion, which can take on the form of what the researcher called a 'Voodoo cult' and likened it to religious rituals, which might actually not even be that bad, something like a cult of 'good science' (see Feynman, 1974 for cargo cult science). This seems not to be some explicated or codified guidance for 'good science' or research integrity, but a lived ethos concerning

conducting research that one is socialised into (Valkenburg et al., 2021; Vincent & Hooper, 2025):

> It somehow needs both and the interface between the two. […] I think many mechanisms are like this Voodoo cult, but maybe even Voodoo cult isn't that bad. […] It certainly belongs to the concept of integrity. […] the personal is only one facet. An important one, but by no means the only one. There are others that are at least as important. The interpersonal is, the institutional and the interaction between institution and person. […] it has to be a social undertaking, in a certain sense. And social in the sense of people among themselves, but also social; the elements are institutions, where it's no longer about the individual person or the people in the group, but rather the spirit that is embodied by them. […] at some point, they hover over us like the Ten Commandments[24] (Professor Biopsychology, Germany).

[24] Original in German: ‚Aber eigentlich ist es mehr oder mindestens genauso viel eben eine Aufgabe der Integrität der Forschung, der Einrichtung. Wenn man es so findet. Also sozusagen des Kollektivs an Subjekten, aber auch der Institutionen, des Kollektivs der Institutionen, die beteiligt sind. Wenn das nur eine einzelne Person wäre. Also man braucht immer zumindest dann die Philosophie oder die Religion, die sozusagen mithält. Das tut es jetzt bei der Integrität im Rahmen von Forschung nicht, sondern da braucht es eben andere überpersonelle was auch immer was. Also Regulare sind ein bisschen Ausdruck dessen, was eigentlich man damit anstrebt sozusagen. Mit integrer und rechtschaffener und reliabler und so weiter Forschung. Und der Aspekt und das wurde glaube ich lange nicht so gesehen, sondern das ist Privatsache. Also es gibt halt Lügner die Lügen und die fälschen Daten und es gibt dann den ehrlichen, rechtschaffenen, integren Menschen, der macht das nicht. Das ist aber natürlich überhaupt, also vielleicht ist es so, aber vermutlich nicht, sondern man sieht unter den Polizisten gibt es genauso viel Diebe wie unter der Normalbevölkerung und unter den Lungenärzten, die Raucherlungen operieren, genauso viel Raucher wie unter den Normalen. Also sozusagen das ist eigentlich ein falscher Maßstab in gewisser Weise, weil es nur individualisierte Verantwortung als einzelne gibt. Das braucht also eben Perspektiven und Initiativen, die eben über Personal und auch über Kollektiv vielleicht das Ganze betrachten und entwickeln. Man muss aber aufpassen, dass es nicht völlig auseinander geht. Also wenn es nur so was geht, da jetzt vielleicht gar nicht mehr angesprochen wird, also die Organisation gar nicht mehr sozusagen ihren Werten wahrgenommen und dann im Alltag sozusagen im operativen Geschäft auch angewendet wird bei den Individuen oder Personen, dann stimmt es auch nicht. Also es braucht irgendwie sozusagen beides und die Schnittstellen zwischen beiden. Und ja, das ist nicht so einfach. Und da gibt es viele Mechanismen. Ich glaube viele Mechanismen sind wie der Voodoo Kult. Aber vielleicht ist auch Voodoo Kult nicht so schlecht. Wenn der Priester da sein Weihrauch schwenkt, ist auch Kult, aber uns eigentlich völlig banal, was soll das bringen? Das bringt doch was. Also es hat in irgendeiner Weise vermutlich eine Wirkung. Das könnte man mit der Wissenschaft untersuchen, kulturwissenschaftlich oder wie auch immer. Vielleicht noch wenigstens naturwissenschaftlich, vielleicht da auch. Aber das ist vielleicht nicht so, das gehört sicherlich dazu zu dem Begriff der Integrität. Wenn man das denkt. Also ich sehe es ja auch stark persönlich, aber das Persönliche ist nur eine Facette. Eine wichtige, aber auf keinen Fall die einzige. Einige andere die mindestens genauso wichtig sind. Das Interpersonale ist das Institutionelle und das

The integrity of the research

Considering that most of the elements listed at the beginning of this chapter as relevant to, components of, or preconditions for research integrity seem to be about the conduct of research or research quality (Helgesson & Bülow, 2023), in this section, I will go through some of the most prominent themes emerging in conversations about the integrity of the research. One scholar succinctly defined research integrity as "an evaluation and assessment of research conduct" (PhD Personality Psychology, Germany). While cases of misconduct are discussed the main theme defining the integrity of the research for the researchers can best be captured by forms of methodological rigour, depending on their knowledge production modes (Ulpts & Schneider, 2025). Hence, impressions of methodological rigour are presented as parts of researchers' credibility management (Daston, 1995; Geiger, 2021; Latour & Woolgar, 1979; Shapin, 1995a).

Misconduct: fabrication, falsification, and plagiarism.

In line with its historical origin as a regulatory concept and policy object (see e.g., Davies & Lindvig, 2021; Guston, 2000; Hesselmann & Reinhart, 2024; Price, 2013; Roberts et al., 2020), as well as the rather punitive and inverse way to capture research integrity in policy documents and research on research integrity (Horbach & Halffman, 2017; Horbach et al., 2024), researchers, as already illustrated, sometimes link research integrity also directly to the issue of research misconduct in the form of fabrication, falsification and plagiarism. The following two quotes focus on misconduct related to the engagement with the data, which is in line with the previously mentioned aspect

---

Interagierende zwischen Institution und Person. […] Also das ist glaube ich sehr ähnlich. Man kann da jetzt viele dieser Aspekte, die wir besprochen haben, sind eben, die sind ja recht falsch, die ausschließlich in der Person verantwortlich zu machen. Aber es wäre unmöglich, die Person da rauszunehmen. Ja, sondern es muss natürlich ein soziales Unterfangen sein, in gewisser Weise. Und zwar sozial im Sinne von Personen untereinander, aber auch sozial, die Elemente sind eben auch Institutionen, wo es gar nicht mehr um die einzelne Person geht oder die Personen der Gruppen, sondern den Geist, der dadurch verkörpert wird. Und auch natürlich die Interessen, auch im Gegensatz dann wieder zu anderen Institutionen oder auch die Gemeinsamkeiten mit den Institutionen zusammen, mit einer demokratischen Grundhaltung und mit einer, keine Ahnung, mit einem verantwortlichen Umgang mit Ressourcen. Also nicht im Sinne von Arbeitszeit des Forschenden, sondern auch, im Sinne was kostet das jetzt in Euro oder auch sonst vielleicht investieren und so weiter. Da gibt es viele Aspekte die da eben deutlich überpersonal sind auch wenn die von Personen ausgeheckt werden und dann im Diskurs entwickelt werden aber irgendwann schweben die über uns wie die 10 Gebote.'

of research integrity, as going where the data takes you and taking the data seriously.

> That is obvious that one does not manipulate their data or fabricate it or something like that. […] I also think no one is doing that here[25] (Postdoc Neuroimaging Lab, Germany).

In the second quote, the issue of falsification and fabrication is also linked to the previously mentioned perceived publish or perish culture in academia:

> Research integrity, that is data falsification, that one again and again fabricates data or to bend the data to fit one's hypotheses to produce quicker and more papers[26] (Lab Leader Biopsychology, Germany).

Furthermore, it is mentioned that 'no one is doing that here'. This is a recurring theme in the talk about deviant research behaviour among interlocutors. The exposure to scientists misbehaving or violating research integrity is usually presented as second-hand accounts and not from within one's own immediate environment of someone's lab group. It seems to be usually about stories that they have heard about what is going on in other labs. This is somewhat in line with a recent report from Feenstra and colleagues (2025) comparing perceived proliferation of research misconduct in one's own university versus all of Spain. Where around 72% of participants indicated a perception of proliferation of at least one type of misconduct in Spain, and around 50% reported perceiving that in their own university (Feenstra et al., 2025).

In the following quote a researcher does not talk about what good or bad scientific conduct is or looks like but directly refers to a person associated with a widely publicised recent case of research misconduct, similar to the Stapel case in psychology about a decade earlier (Derksen, 2021; Larregue, 2024). This might indicate that due to processes of scandalisation sometimes the case itself and the name of the (alleged) fraudster become synonymous with scientific misbehaviour, instead of the actual practices that are at the foundation of the research misconduct at hand (Horbach et al., 2024).

> It's what Francesca Gino did not have (Postdoc Cognitive Science, The UK).

[25] Original in German: ‚Ja genau und natürlich in den härtesten Fällen aber das ist klar dass man nicht seine Daten manipuliert oder erfindet oder so. Das ist natürlich völlig das ist natürlich klar. Aber Ich glaube das macht hier auch keiner.'

[26] Original in German: ‚Äh, Forschungsintegrität, äh, also, also, das ist, ist, ist Datenfälschung, äh, dass man immer wieder Daten komplett erfindet oder aber eben sie so hinbiegt, dass sie, äh, zu seinen Hypothesen passen, um schneller oder, äh, mehr Paper zu produzieren.'

Yet another researcher recalls an incident as a reviewer where a manuscript that had to be evaluated seemed to contain copy-pasted passages from other publications, which the researcher called out as plagiarism and described notifying the editor. However, the scholar also mentioned being dissatisfied with the journal's handling of this case of apparent misconduct, reducing the value the journal has for the researcher afterwards, because there is an expectation of a gatekeeping mechanism provided by the editors of scientific journals and their responsibility to ensure the integrity of the scientific record (as was previously stated by the PI of a Dutch neuroimaging lab).

> I once had a manuscript there even, in my opinion, was plagiarism. They partly just copy-pasted sentences from other studies. There I simply wrote to the editor and said: sorry, but I would not review that at all. To me that was absurd, that that was even sent out to peer review. Did really not speak for the journal at all, in my opinion[27] (Postdoc Neuroimaging Lab, Germany).

Methodological rigour

When talking about where notions of research integrity or good scientific conduct come from, researchers sometimes talked about methods courses, indicating the centrality of methods in their research (Hangel & Schickore, 2017; Schickore & Hangel, 2019). During the conversation with a postdoc from the German neuroimaging lab, the practice of peer review came up, and I asked what the postdoc focused on when reviewing manuscripts of other researchers, to grasp what the postdoc sees as important and relevant to assess research as a proxy for what 'good science' or research integrity contains for them (Laudel, 2024). The postdoc talks about the importance of methodological rigour and how that relates to how credible they consider that research to be:

> The methodological rigour, […] For me it is about, okay, how well have they done it? So, how much can I trust that what they have done, and how they then have analysed the data?[28] (Postdoc Neuroimaging Lab, Germany).

---

[27] Original in German: ‚ich hatte mal ein so ein Manuskript, da habe ich dann einfach, also tatsächlich waren da, also da wurde sogar, also meiner Meinung nach waren da sogar so Plagiatsfälle drin, die hatten teilweise aus anderen Studien einfach so Sätze so reinkopiert, und also da habe ich einfach dem Editor geschrieben und meinte so, sorry, also das würde ich überhaupt gar nicht erst reviewen, ich fand es auch absurd, dass das rausging zum Peer Review, hatte jetzt auch echt nicht für das Journal gesprochen meiner Meinung nach, ja, genau, da habe ich einfach nur halt dem Editor geschrieben und meinte, das geht einfach nicht, das ist, kann einfach gar nicht, ist überhaupt nicht einem Peer Review würdig.'

[28] Original in German: ‚das ist auf jeden Fall so die methodische Rigor, wie sagt man das, Rigorosität, glaube ich, ja, genau, also, dass ich, also, wenn ich reviewe, also, ich reviewe

Hypothesis testing

One of the themes that reemerges concerning their research methods is the centrality of hypothesis testing in their conduct (Schickore & Hangel, 2019). Representative of this high valuation of hypothesis-driven research is the following remark of a researcher who talks about the importance of doing hypothesis-based research and knowing what an a priori hypothesis is, calling them 'the basics of good research practice'. The recognition, value, and protection of confirmatory research is a topic that will be explored further but also complexified in Chapter 7.

> There was this course called research methods. And there we, of course, learned that one should conduct hypothesis-based research. […] What is an a priori hypothesis? […] those are the basics of good research practice[29] (PI Neuroimaging Lab, Germany)

Research ethics ('do no harm')

Another issue that came up when trying to imagine what research integrity might mean to the researchers was the issue of research ethics. More precisely, it revolved around the crucial aspects of how-to treat participants in the form of doing no harm and what kind of relationship to have with them, similar to the qualitative research concept of building rapport (Duncombe & Jessop, 2002). Next to how to 'treat everybody in your research, for example research groups, with respect and also be honest in the relationships' (Postdoc Neuroimaging Lab 2, The Netherlands). There is also talk about the bureaucratic procedure of institutional ethics approval as providing a protocol to follow in the execution of the research (ethical issues will be further explored in the next chapter on research ethics in the lab). As the following quote illustrates (Shaw & Satalkar, 2018):

> That one, for example, also writes an ethics application and just states here this is what I am going to do and then not deviates from that (PhD Biopsychology, Germany).

---

nicht mit der Idee, wie neu ist das, was sie da gemacht haben, sondern ich reviewe mit der Idee, okay, die haben dieses, diese Studie gemacht, die wurde ja jetzt auch schon gemacht, also, ist ja jetzt auch totaler Quatsch, im Nachhinein zu sagen, finde ich total langweilig, und dann geht es mir darum, okay, wie gut haben die das gemacht, also, wie sehr kann ich dem vertrauen, was sie da gemacht haben, und wie sie die Daten dann auch ausgewertet haben.'

[29] Original in German: ‚Ich habe ja Psychologie studiert und da gab es als, im Studium gab es ein Fach, das hieß Forschungsmethoden. Und da hat man natürlich schon gelernt, man sollte Hypothesen basiert forschen, also was uns immer so selbstverständlich erscheint, was glaube ich in der Medizin häufig eher nicht so ist. Was ist eine a priori Hypothese? Also das geht ja schon in, das sind ja die Grundzüge guter wissenschaftlicher Praxis'.

However, at the same time, as previous research already suggests, researchers do also often see the requirements and bureaucratic demands from institutional review boards nearly as an annoyance and an additional burden to doing their research (De Vries et al., 2006; Elizondo & Kaltenbrunner, 2024; Pickersgill, 2012). During my field observations, I was able to follow a meeting, at the neuroimaging lab 2 in the Netherlands, discussing a large-scale research project with many cooperating labs, where the researchers were discussing preparing for and getting ethics approval, as the following excerpt from my field notes illustrates:

> The researchers discuss going through and preparing as well as filling in the ethics protocol. It was then quickly mentioned that they could take a previous, already approved ethics protocol and adapt it accordingly to reduce the workload of having to go through the whole procedure again.

Such frustration over the way ethics protocols can be institutionalised is exemplary demonstrated by the following statement:

> There are some IRBs, especially in the US, that are a little bit overzealous (Reader Cognitive Science, UK).

Prevention of bias

The next theme concerning the integrity of research is the prevention of bias. This seems to be connected to a notion of purity when it comes to doing the work as a scientist, similar to ideals of value-free science or notions of disinterestedness (Merton, 1942). Somewhat in congruence with the tendency of the guidance documents for research integrity, which also show some tendencies towards the value-free ideal of science (Ambrosj et al., 2023), this is described as being driven by the aim to gain knowledge without being distracted by anything else. Practically conceptualising research integrity and 'good science' as the purity of research, clean from non-scientific, social and human disturbances that might intervene in the production of objective scientific knowledge (Daston, 1995; Daston & Galison, 1992; Field & Derksen, 2021; Flis, 2019).

> You're driven to do your science, because you want to know something. So, it's driven towards gaining knowledge and that is the only thing it should be about [...] if you're focused on that drive solely, you will also not be distracted by other motives (PhD Neuroimaging Lab 1, The Netherlands).

One scholar talks about laying a 'veil of ignorance' between the research and the researcher, similar to the practice of blinding, as a layer of protection against bias and motivated reasoning. In line with Daston and Galison's (1992) notion of mechanical or noninterventionist objectivity as mediums of

credibility, one might see this as an attempt to keep scientific practice pure, neutral, objective, innocent and clean to convey credibility. This is reminiscent of a lot of the motivation behind the Open Science reform movement, which seeks to protect science and research integrity from human bias (see e.g., Flis, 2019; Morawski, 2020). As Flis (2019) and Morawski (2020) argue, this recent science reforms' focus on protecting the scientific processes from human bias and motivated reasoning might be a case of the domain-specific scientific knowledge gained in psychology and the wider cognitive sciences, about cognitive biases and the limits of rationality in human cognition, being applied to science itself, something like a new psychology of science (Aczel, 2023).

> We're all hopelessly motivated reasoners and confirmation biased people. Constantly renewing your mantra that you are the worst judge as to whether you can or can't and then trying to veil of ignorance over yourself in how you take the next step. But I think actually in the long run that's the only way to not get yourself stuck in a hole where you have made the wrong conclusion (PI Cognitive Science, The UK).

Societal relevance and practical utility

However, this purity and freedom of scientific knowledge production seem to others not the only thing that matters. Some also talk about the importance of the societal value and practical usefulness of their research. Such social responsibilities of science naturally touch on the social contract between science (scientists) and society (Douglas & Branch, 2024), and how responsible conduct of research or research integrity relate to it (Pols et al., 2024). This is also part of an established discussion in research governance and research evaluation concerning how independent research can and should be from societal and cultural issues and what the role of academic freedom is in all of this (Kulczycki, 2023; Wolfe, 2018).[30]

> I would like it to be more about whether we conduct useful research [...] Is the work we do here actually useful to society in some way? [...] ultimately, we're supposed to do society something good. And I think, thinking about whether what you do does that [...] should be part of research integrity (PhD Neuroimaging Lab 1, The Netherlands).

While the interlocutors state the conviction that demonstrating research integrity is an ethical obligation to deliver 'good science' as promised, an intriguing bit is the following statement by one researcher that turns this perceived

[30] With a historical perspective, part of that discussion are also investigations into how scientific objectivity was and is not just a scientific concern, but also part of political strategies (Wolfe, 2018).

responsibility of science towards the wider society directly onto science itself. Accordingly, the scholar states that research has integrity if it is conducted in a way resembling laypersons' understanding of science. Adding that laypeople think science is done very systematically and self-critically, while in reality, that is not always the case.

> The ethical obligation to deliver the knowledge that you are supposed to deliver. It's kind of like you have integrity if you conduct your science in a way like laypersons think you are conducting it. […], I think doing science as in the ideal way, so really very systematically, very self-critically. I think this is integrity. Striving to do both transparent and correct methods based on the current highest standards (PhD Personality Psychology, Germany).

The relationship between research integrity and Open Science

Due to the previously mentioned often assumed close or even mutually reinforcing relationship between Open Science and research integrity (see e.g., Bouter, 2024; Haven et al., 2022), the prominence of Open Science related topics in the current research integrity literature (see e.g., Resnik & Shamoo, 2017; Sijtsma et al., 2021; Tang, 2025b), Open Science having found its way into research policy documents around the globe in various forms (Labib et al., 2025; Laine, 2018; Manco, 2022; UKRIO, 2025), the fact that some of the scholars that I did my fieldwork with were actively involved in Open Science, and Open Science themes like reproducibility, preregistration, transparency, and forms of being open occurring throughout the conversations, in this section I will briefly explore how the researchers perceive the relationship between Open Science and research integrity.

> So, I care a lot about open science and reproducible science, and I think I would like research integrity to mean those things (Postdoc Neuroimaging Lab 1, The Netherlands).

Considering the ongoing Open Science reform movement, especially in the social, cognitive, and behavioural sciences, with psychology and biomedicine as the epicentre (Bartscherer & Reinhart, 2025; Peterson & Panofsky, 2023), an important question remains just how the concept of research integrity relates to Open Science (Haven et al., 2022). Some of the remarks from scholars during the interviews indicate a close relationship between Open Science and research integrity. The next quote illustrates a researcher's position regarding Open Science as a type of science that can also work against the perceived publish or perish culture, or as the scholar calls it, the 'subjugation of science to publishing'. The PhD talks about how Open Science practices can enable

accountability through opening up one's work to the 'scrutiny' of others and are seen as encompassing 'good practices' or even the 'only way' towards 'better science':

> The principle that you have to abide by is, it's got to be openness and publicness. And this is why we need at the same time to pull out of this publish or perish culture and this total subjugation of science to publishing in some limited amount of papers and journals. We need to emancipate from, it's only relevant if you publish it in a reputable paper. At the same time, we need to go with the open science initiatives. It's the only way (PhD Neuroimaging Lab, Germany).

However, it is also suggested that Open Science practices can be a hindrance, because they do require more work like needing data curation, adding to the already heavy academic workload (Hostler, 2023). Furthermore, the response from the researcher suggests that, while the potential for accountability might be provided, the suspicion is that a vast number of studies will be stored on some repository practically becoming Open Science's own graveyard of accessible but unchecked studies, like Ferguson and Heene's (2012) 'vast graveyard of undead theories' (p. 555):

> Still, it's not because you publish all of your preliminary results or whatever that someone is going to replicate your study. There will be still a lot of studies that nobody cares about, and they will not get replicated, they will not get scrutiny, they will just hang out there in the vast black that is the internet. But at least they are there and there's accountability. The problem is even good practices, even open science, again it takes, it can be complicated. Sometimes it's extra steps in your way to just get on with your research. You need to organise your data you need to publish them, you need to put them in a GitHub, make them available (PhD Neuroimaging Lab, Germany).

Again, indicating the perceived importance and close link between notions of research integrity and how to treat the data and linking it to Open Science, one researcher directly mentions the FAIR (Findable, Accessible, Interoperable, and Reusable) principles (Wilkinson et al., 2016). Strikingly, the scholar indicates that there might be a trade-off between how detailed or transparent one reports their research and how informative such reports are for others. The scholar, while wanting to have code and data shared as much as possible to, among other things, enhance reproducibility, still mentions an apparent dilemma between these Open Science practices and research ethics. Here, the ethical issue is the sensitivity of the data on which the research is based (Moors, 2026). Hence, if Open Science and research integrity are equated or Open Science is seen as a crucial component of research integrity, then there might be critical conflicts between research integrity and research ethics (see e.g., Moors, 2026; Prosser et al., 2024; van Drimmelen et al., 2025).

> The FAIR principles of data come to mind [...] that it's clear where something comes from, what this entails, to be as detailed as possible while still keeping it informative for somebody who hasn't done the research. [...] Of course, if you're working with sensitive data, it might not be, but I think every data set should be published at some point (PhD Neuroimaging Lab 2, The Netherlands).

When talking about an Open Science course taught by the postdoc, I was able to observe, the researcher also elaborated on the types of procedural interventions aimed at improving science, which the postdoc wants the students to support and implement, like preregistration and registered reports (Penders, 2022). However, at the same time, the researcher does not want them to practice something like mindless Open Science by just following some checklists (see Gigerenzer, 2004, for *mindless statistics*). According to the postdoc, one has to be reflective about these things, because checklists are 'hackable' and the research behind open data and preregistrations can still be 'rubbish'.

> I really don't want to give them some kind of checklist to use to make assessments about what is good, because checklists are also hackable. So, many papers in psychology are now pre-registered or have open data and such, but they're still rubbish. So, [...], just think about it, reflect on it[31] (Postdoc Personality Psychology, Germany).

When I directly inquired another German postdoc about whether Open Science and research integrity are the same or how these two notions relate to each other, a somewhat more complicated picture emerged; the interlocutor stated that 'It depends a bit on what all counts as Open Science' (Postdoc Personality Psychology, Germany). While some practices and standards, like reproducibility, that can fall under the hypernym Open Science are seen as requirements for research integrity and quality, others are just seen as a means to an end, and it is stated that research can still be of high quality without using Open Science practices.

---

[31] Original in German: ‚dass man einfach so merkt, so, nee, Alter, das stimmt nicht und ich wurde verarscht und dann irgendwie auch ein bisschen dann die Energie entfaltet, skeptisch zu sein in Zukunft irgendwie. Ja, und dann will ich ihnen eigentlich am liebsten nicht irgend so eine Checkliste an die Hand geben, um gut zu beurteilen, weil Checklisten sind halt irgendwie auch hackbar. Also viele Paper in der Psychologie sind jetzt präregistriert oder haben offene Daten und so, die sind trotzdem Schrott. Also eher, ja, halt denken, darüber nachdenken. Es ist oft auch eine Klausuraufgabe von mir, dass ich ihnen eine Studie gebe, die halt irgendwas macht, aber nicht gut und dann sollen sie sich eine bessere überlegen. So. Das hätte ich gerne, dass sie konstruktiv Fehler feststellen und sich bessere Studien vorstellen können und nicht schnell sagen, so, ja gut, aber das ist jetzt halt das, was wir haben, damit müssen wir jetzt irgendwie auch zufrieden sein und dann ist das auch die Antwort auf meine Frage. Bestmögliche Antwort. Aber klar, also ich meine, ich will sie schon für Registered Reports begeistern und für Transparenz.'

> The question is whether reproducibility is part of it. Thus, when analyses aren't reproducible, […] then that is somehow completely invalid. So, there are certain things that are necessary. However, now open data or something like that, there you can of course do completely valid good research, without sharing your data. It is just that when you don't share your data, then it is hard to verify whether the analyses are reproducible and robust and so on. Therefore, I think Open Science is in principle a means to the end of doing better research. When it comes to open data and open code, they're mostly also a tool to verify that the research meets quality standards (Postdoc Personality Psychology, Germany).

Hence, the researcher highlights that Open Science contains multitudes, and it might be problematic to talk about it as just one unitary thing when relating it to research integrity (Fecher & Friesike, 2014; Leonelli, 2023; Bazzoli, 2022). This further complexifies the relationship between Open Science and research integrity, since research integrity itself is seen as multifaceted and dependent on the specific research and context (Penders et al., 2009). Hence, the relationship between Open Science and research integrity most likely varies between research contexts. Considering the aforementioned component of research integrity of being willing to put yourself out there and be prepared to be held accountable for one's research, Open Science is ultimately described as a means to that end (Bouter, 2024; Haven et al., 2022). According to the scholar, without Open Science practices like open data and open code, others could not check the integrity of the research. Furthermore, touching on the responsibility of journals by referring to changes in the editorial and peer review procedures at the journal *Psychological Science* following the appointment of prominent Open Science reformer Simine Vazire as editor in chief (Vazire, 2024), the researcher also talks about how requirements and gatekeeping processes in accordance with Open Science could lead researchers to check already and improve their research themselves (Nosek et al., 2015). Moreover, preregistration is seen as a procedural intervention with a direct causal effect on research quality by preventing potentially problematic practices and enabling transparency (Nosek et al., 2018).

> The moment you require that, the people naturally have a reason to take the initiative. So, the moment PsychScience is now saying we check the analyses when accepted, then before you survive the whole peer review process and are rejected in the end, because it is irreproducible, you yourself check it again. And that is, I think with preregistration and registered reports different again. This is really a more direct causal effect on the quality, by preventing certain practices,

but also the researchers create a commitment that they don't do that if possible[32] (Postdoc Personality Psychology, Germany).

However, then I also asked whether there might be situations where conflicts between Open Science and responsible conduct or research integrity could arise, to which the researcher replied:

Scenarios where you only somehow optimise for this one thing and then, for example, only collect data that can be shared without any problems and, therefore, conduct no more qualitative studies, because it is more complicated to share the data. I would see that as a potential conflict. [...] I do actually see that when we say Open Science in a broader sense also concerning larger samples and so forth. I see that in the USA something like that always happens quicker, this ruthlessly optimising. Okay, now we must produce larger samples, from now on we only use MTurk. And I do belief that this actually happens, but it mainly happens in these hyper-competitive high-pressure scenarios[33] (Postdoc Personality Psychology, Germany)

---

[32] Original in German: ‚Ja, also es kommt ein bisschen darauf an, was man alles unter Open Science zählt. Aber jetzt also so zum Beispiel ist die Frage, ob Reproduzierbarkeit drunter fällt. Also wenn Analysen nicht reproduzierbar sind, aus meiner Sicht, dann ist es irgendwie komplett hinfällig. Also da gibt es bestimmte Dinge, die notwendig sind. Aber jetzt zum Beispiel Open Data oder so. Da kannst du natürlich vollkommen valide gute Forschung machen, ohne deine Daten zu teilen. Es ist nur, wenn du die Daten nicht teilst, ist es schwerer zu verifizieren, ob die Analysen reproduzierbar und robust sind und so weiter. Also ich denke Open Science ist im Prinzip was, also ein Mittel zum Zweck bessere Forschung zu machen. Wenn es um Open Data, Open Code geht, vor allem wirklich auch ein Werkzeug, um zu verifizieren, dass die Forschung Qualitätsstandards entspricht. So und dadurch aber natürlich in dem Moment, wo du es einforderst, haben natürlich dann die Leute einen Grund, das initiativ zu machen. Also in dem Moment, wo jetzt PsychScience sagt, wir rechnen die Analysen dann nach, wenn sie akzeptiert sind. Dann bevor du den ganzen Peer Review Prozess überlebst und am Ende rejected wirst, weil es sich nicht reproduzieren lässt, checkst du es natürlich selber noch mal. So und das ist in der Hinsicht. Ich glaube bei Pre-Registration, Registered Reports ist es nochmal anders. Das ist wirklich ein direkterer, also direkterer kausaler Effekt auf die Qualität, in dem man man bestimmte Praktiken unterbindet, aber auch die Forschenden schaffen sich ein Commitment, dass sie das möglichst nicht machen. Allein dadurch, dass man es preregistriert hat und es damit transparent macht, unabhängig davon, ob es dann gecheckt wird oder nicht. So, aber es ist keine simple Beziehung. Und ich glaube auch schon, dass Forschungstransparenz schon auch unabhängig ein Wert ist. Also ich mache viel Open Science als Mittel für bessere Forschung und so weiter. Aber ich glaube schon, dass es auch unabhängig davon tatsächlich einfach mit der Rolle der Wissenschaft in der Gesellschaft dass es auch unabhängig davon. Also wenn es jetzt eine schlechte Studie ist, und die wird so oder so publiziert, dann lieber mit offenen Daten.‘

[33] Original in German: ‚Also, das wäre jetzt, das sind immer diese Szenarien, dass man jetzt nur noch auf das eine irgendwie optimiert und beispielsweise jetzt nur noch Daten sammelt, die man problemlos teilen kann und deswegen keine qualitativen Studien mehr macht, sodass es komplizierter ist, die Daten zu teilen. Das würde ich als potenziellen Konflikt sehen.

The postdoc considers a process of hyper-optimisation in conjunction with a focus on just one thing like data sharing as a potential conflict, because it could result in lowering the amount of qualitative research being done, since such data is harder or sometimes even impossible to share (which is a widely discussed issue in the confrontation of Open Science with qualitative research and epistemic diversity, see e.g., Cole et al., 2024; Malich & Rehmann-Sutter, 2022; Prosser et al., 2024). Another issue pointed out is the Open Science movement's recommendation or demand to increase sample size to increase statistical power, which has seemingly led to the use of MTurk. Which can also be problematic due to the concerns about the data quality from MTurk (Chmielewski & Kucker, 2020). This is referred to as ruthlessly optimising, because it seems to be seen as a form of tunnel vision, where people ignore complexity, complications and nuance in the application and conceptualisation of Open Science in the context of a diverse and complex reality affecting the relevance and feasibility of Open Science, as well as trade-off concerning other components of 'good research' (see e.g., Leonelli, 2023; Ulpts & Schneider, 2025).

Similar to the earlier-mentioned sentiment of Open Science being a means to an end, the PhD involved in metascience clarifies that well done transparency and Open Science practices are necessary for 'good science', but they are not sufficient. They are required to be able to know whether the research itself is good or bad. The PhD outlines the crucial difference between transparency (Open Science) and rigour (research integrity), which is mostly the correctness of methods, again highlighting the previously outlined perceived importance of the application of methods and methodological rigour for the integrity of the research. According to the scholar Open Science and research integrity are 'definitely not the same'. Then the researcher considers whether the best way to conceptualise the two is that the goodness of science or rigour is the overarching thing, while transparency (Open Science) is one of the necessary

---

Es ist immer die Frage, ob das wirklich so die Leute so auf die Anreize in der Form anspringen weil es ja immer mehrere Anreize gibt und Trägheit in dem System und so weiter ich sehe schon dass also nur bei den jetzt wenn wir jetzt Open Science im weitesten sinne auch jetzt hier größere Samples und so weiter also ich sehe das immer in den USA passiert so was immer schneller aus meiner Sicht das hat dieses ruthlessly optimizing okay jetzt müssen wir große Samples machen jetzt machen wir nur noch m-Turk. Und ich glaube schon das passiert schon das passiert vor allem aber in diesen hyper kompetitiven high pressure Szenario ich habe das Gefühl und ich weiß nicht ob das einfach nur Europa mehr friction gibt im System oder träger auch vielleicht einfach nicht so innovativ dass man sofort aufspringt ich wüsste zum Beispiel gerne wo wie viel Covid-forschung rausgebracht wurde weil gefühlt war das sehr viel USA und China das ist meiner Sicht die in [...] meisten vielstrebige Optimierung des Publikationsoutputs des Outputs betrieben wird ich weiß aber nicht was das ist was das hier.'

conditions and together with the correctness of the methods, they're sufficient for research integrity.

> No, they're definitely not the same. […] Maybe the best terminology is to say that rigour or goodness of science is the overarching thing. Transparency is a part of it. It's necessary for rigour or goodness, but there's also the correctness of the methods. That is also another part that is also necessary. So, both together, they're sufficient for rigour (PhD Personality Psychology, Germany).

Related to some of the comments about transparency and Open Science is a more general notion that can also exist independent of Open Science, which was described by German respondents with the German word 'Nachvollziehbarkeit'. Similar to the German word 'Nachprüfbarkeit', this does not have a direct translation in English (Ulpts & Schneider, 2025). 'Nachvollziebarkeit' is somewhat of a combination of being able to trace and understand what was done or happened:

> And then I think it is important to see, okay how do they now get to their results, so what have they done in their analysis and can I understand/ trace (nachvollziehen) that. […] Well, I always have the feeling, okay, if I can't understand it 100%, then sorry, then I can't say that I think it's okay[34] (Postdoc Neuroimaging Lab, Germany).

In a similar vein, another researcher linked this aspect of 'Nachvollziehbarkeit' to the scientificity of the research (how scientific it is), because it makes the researcher and the research vulnerable, which is again in line with the previously mentioned element of research integrity of owning up to your research and being willing to be held accountable for your conduct as a researcher. Such statements could be interpreted as research integrity requiring 'putting popper to work' and presenting conjectures in a way that implies the risk of, if not falsification (Derksen, 2019, p. 449; Popper, 1959), then critical appraisal.

> In general, I think that in research it is something that somehow gives science a certain scientificity and also comprehensibility/traceability (nachvollziehbar) in

[34] Original in German: ‚Und dann finde ich es halt wichtig zu schauen, okay, wie kommen sie jetzt zu ihren Ergebnissen, also, was haben sie in der Analyse gemacht, und wie gesagt, kann ich das, kann ich das nachvollziehen, also, ich finde es, also, der schlimmste Fall, was ich, glaube ich, schon mal hatte, war, wo ich so ein Review, also, ein Manuskript gereviewt habe, und es war absolut unmöglich, die Schritte nachzuvollziehen, und wir haben einfach, keine Ahnung, wie viele Runden gehabt, vier Runden oder so, weil die Autoren einfach immer wieder nicht in der Lage waren, mir so zu erklären, was sie genau gemacht haben, dass ich es einfach verstehe. Also, ich habe halt immer das Gefühl, okay, wenn ich es nicht 100 Prozent verstehen kann, dann sorry, dann kann ich das halt nicht sagen, dass ich das so okay finde, irgendwie, und ich fand es super anstrengend, weil die immer so, nur so Brocken hingeworfen haben, und das war aber nie so ganz klar für mich, also, ja.'

> the simplest sense. […] there is also the fact that science should be comprehensible/traceable and therefore also vulnerable (angreifbar). Vulnerable (angreifbar) – that is now such a negative word, but it should be assessable[35] (PhD Personality Psychology, Germany).

This also aligns with the sentiment that Open Science is just a means to an end; in this case, it seems to be one way to achieve the more general epistemic aim of 'Nachvollziehbarkeit' or 'Nachprüfbarkeit'. This begs the question of what happens when the means become the end? In other words, considering that Open Science and research integrity are not actually seen as the same, what are the consequences when Open Science practices like preregistration or registered reports are used as indicators for credible and 'good science', instead of a window that provides access for the evaluation of credibility and quality (more on this also follows in the preregistration chapter (see e.g., Devezer & Penders, 2023; Klonsky, 2025; Penders, 2022)?

Discussion and conclusion

In this chapter, I presented a thematic analysis of interviews and fieldnotes based on my ethnographic fieldwork in psychology, neuroscience and cognitive science labs in Germany, the UK, and the Netherlands to capture 'integrity talk' and answer the overarching research question of how researchers understand and enact research integrity or 'good research' in their everyday activities (Davies, 2019).

While the researchers are unfamiliar with the phrase research integrity, and it feels foreign to them, once they begin to reflect on what it might mean in the context of their work, a multifaceted, complex and somewhat contradictory picture begins to emerge (Shaw & Satalkar, 2018). The lack of familiarity with the concept or even term research integrity is in line with previous research indicating a disconnect or areas of misalignment between research integrity policy and the experiences and values of actual researchers (Brenninkmeijer & Pols, 2026; Davies, 2019; Davies & Lindvig, 2021; Hangel & Schic-

[35] Original in German: ‚Allgemein in der Forschung ist es glaube ich etwas, was irgendwie der Wissenschaftlichkeit auch so eine Nachvollziehbarkeit gibt auf (in) dem einfachsten Sinne. Also Ich sage, dass für mich immer so das ist, ja kein Hexenwerk was man macht ich mache ja nix Geheimes und warum sollten jetzt zum Beispiel Analyse Skripte, die ich geschrieben habe, die eh im Artikel stehen warum sollen die nicht öffentlich zuganglich sein? Also das ist so dieser ganz grundsätzliche Punkt das ist ja nichts es gibt ja nichts das ich ein finanzielles Interesse daran hätte oder so. Auf der anderen Seite das ist ja auch, dass die Wissenschaft nachvollziehbar und dadurch ja auch angreifbar sein soll. Also angreifbar das ist jetzt so negativ gewählt aber das soll beurteilbar sein und ich finde wenn der Kern also ganz spezifisch jetzt so Analyse Sachen weil ich ja empirisch gearbeitet habe quantitative sind natürlich die Analysen das kernstuck von dem Artikel.'

kore, 2017; Pols et al., 2024; Shaw & Satalkar, 2018; Valkenburg et al., 2021; Vincent & Hooper, 2025). Once researchers talked about 'good research', I categorised their free association into features of the researcher, the research system as well as the group of researchers, and the research (Helgesson & Bülow, 2023). In line with the 'bad apples' framing of deviant conduct (Lambert & Degn, 2026), research integrity was linked to the moral scholar who is conscientious, takes responsibility and ownership for their work, is honest and humble as well as reflexive (Field et al., 2024; Valkenburg et al., 2021). Honesty and humility, as critical reflexivity, could also be conceptualised as parts of research conduct, since they were usually described as forms of reporting, but since they were also tight to the person I present them here as characteristics of the scholar. The role of the research system was often presented in a rather negative light as working against research integrity with perverse incentives and a publish or perish (impact or perish) culture.

The collective or group of researchers was sometimes presented as a largely trust-based system that builds on the mutual dependence of researchers for knowledge production and correction (Whitley, 1984, for mutual dependence, for research integrity and collaboration see Pols et al., 2024; Brenninkmeijer & Pols, 2026). However, at the same time it was also mentioned that to ensure 'good research' it requires a certain degree of distrust and (expected) scrutiny from a competitive community. I distinguish between the research system and the group of researchers as a community, because the first is more about rules, incentives, and institutions, while the other is more about something like an ethos of the collective and scholars' interdependence resulting from the social dynamics of the community (Vincent & Hooper, 2025).

Lastly, we do see utterances that describe research integrity as the opposite of misconduct, but when they actually talk about research integrity, some of the characteristics of the research that come up can be subsumed under specific notions of methodological rigour, like the role of hypotheses testing due to the centrality of hypotheses driven research (Hangel & Schickore, 2017; Schickore & Hangel, 2019), research ethics (more on why this can also be seen as a component of methodological rigour in the next chapter; also see Mickel et al., 2023; Pickersgill, 2012, 2021), and the prevention of bias and protection of objectivity (Daston, 1995; Flis, 2019). Next to methodological rigour some interlocutors also brought up that the research should have practical and societal relevance, which can conflict with some forms of academic freedom and interpretations of objectivity (Wolfe, 2018).

Research quality, 'good science' and research integrity might be somewhat conflated in the responses from the scholars, not just because they themselves struggle to explicate what research integrity means for them (Gulbrandsen, 2000; Hangel & Schickore, 2017), but also because I often asked about 'good

science' or 'good research', since the researchers indicated not knowing the term research integrity and finding it weird. Therefore, the way I asked may also have contributed to this terminological conflation of research integrity, research quality, and 'good science' or 'good research' as well as the focus on the integrity of the research in their responses.

Considering the recent upstream of Open Science related reform movements throughout the science system and across the globe as well as their positioning within the research integrity community and policy landscape, it is important to understand how these two concepts, Open Science and research integrity, relate to each other (Bartscherer & Reinhart, 2025; Bouter, 2024; Haven et al., 2022; Manco, 2022; Peterson & Panofsky, 2023; Tang, 2025a). Open Science is largely seen as a means to an end, to achieve research integrity. However, according to the researchers, it is not just that Open Science and 'good science' are seemingly not the same, and that completely Open Science can still be 'bad science', the procedures that make science open, as any research practice, can also be poorly done. This in combination with the existing literature reflecting on the potential issues inherent to current directions and application of Open Science suggests that when OS procedures are carelessly used as indicators for research integrity or 'good science', then Open Science would instead of fulfilling its promise of improving science, worsen the situation concerning research integrity. Thus, in the socio-epistemic process of appraisal, instead of enabling the assessment of research quality and the conduct behind research outputs, it actually can obfuscate and hide bad conduct behind a 'superficial veneer of rigor' (Devezer et al., 2021, p. 16; Devezer & Penders, 2023). This potential downward spiral of research integrity in the name of improvement through the promised revolution of Open Science can further be escalated through a process of obliteration, where some OS procedures are used as indicators for research quality and, for instance, function as inclusion and exclusion criteria in literature reviews (Devezer & Penders, 2023).

Devezer and Penders (2023) elaborated on precisely that question when linking Open Science reform to citation politics and a 'bureaucracy of oblivion' (p. 857). They take Folk and Dunn's (2023) systematic review as an example, who excluded 89 percent of the relevant literature due to a lack of preregistration, and warn that perfectly valid and 'good research' can be obliterated through a process of denying citation and inclusion in literature reviews for publications that are not preregistered (Devezer & Penders, 2023). Other examples of the corruption of scientific practice due to Open Science practice becoming the end instead of the means, which can ultimately harm instead of improve research integrity, will be discussed in Chapter 7 on preregistration.

Due to the inherent interconnectedness between research integrity and research ethics, especially in the human and medical sciences (Kolstoe & Pugh, 2024; Pickersgill, 2012) and the prominent role the theme played during my field observations and the interviews, in the next chapter I will go further into detail about how, from the perspectives of the researchers, research ethics is enacted in the lab and relates to 'good research'. Furthermore, in face of the prevalence of Open Science and themes like preregistration and replication as well as reproducibility as a component of methodological rigour in the last two chapters I will explore how the researchers associate the practices of preregistration and replication to research integrity or 'good science'. Hence, the last two chapters provide a further investigation into the researchers' view on the interrelation between research integrity and Open Science as well as a practice perspective on what 'good' or 'credible science' means for them (Pickering, 1992; Valkenburg et al., 2021).

Chapter 6: Ethics in the lab

Extending on some of the themes from the previous chapter, in this chapter I explore the researchers' conceptualisations and experiences with the ethicality of their work as researchers in their laboratories as well as how those relate to their notions of 'good research'. As such I here focus on 'ethics in practice', not on 'procedural ethics' (Guillemin & Gillam, 2004; Potthoff et al., 2023). I first provide some background to situate the discourse about ethical conduct in the context of research integrity and 'good science'. Then I will explore the ethics in practice in the context of their engagement with participants. Subsequently, I present their notions of ethicality in their interactions with colleagues in the lab. Lasty, I conclude on the relationship between the ethicality of conduct and 'good research'.

Introduction

Despite previous attempts at clarifying the relationship between research ethics and research integrity, it remains somewhat unclear how these concepts relate to each other (see e.g., Kolstoe & Pugh, 2024; Shaw, 2019; Steneck, 2006). However, it is also undeniable that, especially when human or animal subjects are involved in the research, they are closely related. Moreover, similar to research integrity guidelines and policies, research ethics standards, also function as instruments for research governance and regulation through bureaucratic and managerial processes inherent to the academic regulatory systems like ethical institutional review boards and standardised demands of ethics approval, for research involving human or animal subjects, from many scientific journals (see e.g., Kolstoe & Pugh, 2024; Moors, 2026; Pickersgill, 2012). The emic perspective on the ethical conduct as rapport with participants and the lab climate as the type of relationships researchers have with their colleagues in their research group can be seen as translation of the policy principle of 'Respect', that can be found in codes of conduct like the European Code of Conduct for Research Integrity (ALLEA, 2023), into practice. Similar to some remarks by the researchers in the previous chapter, some studies indicate that, it often seems to be a dreadful experience for researchers to go through the standardised, formal and bureaucratic process of ethical approval demanded by the institutional academic regulatory bodies (De Vries et al., 2006; Pickersgill, 2012, 2021), or even perceived as disadvantaging certain types of research like anthropological approaches and ethnographies (Herzfeld, 2023; Moors, 2026; Shore, 2024). It also remains questionable whether

these ethics review boards actually fulfil their intended regulatory governance function towards ethical conduct of research (Mickel et al., 2023; Smith-Doerr & Vardi, 2015). Just because a study has received ethic approval does not (necessarily) imply that the resulting publication is based on ethical conduct.

Furthermore, due to 'normative uncertainty and the overdetermination of ethical practice [...], even within the same institution, there may be a variety of coexisting regulatory frameworks constitutive of the regimes of normativity structuring laboratory life' (Pickersgill, 2012, p. 593-595; also see van Drimmelen et al., 2024, 2025; and for similar comments in the context of Open Science see Simm & Eigi-Watkin, 2024). Similar to Daston's (1995) notion of moral economies of science (also see Shapin, 1995a), Pickersgill (2012) describes regimes of normativity as the combination of values, emotions, and norms that can be found in formal and informal discourses and practices which build the social foundation for what counts as 'good science'. Extending on classical lab ethnographies and existing STS work into the interrelations between scientific knowledge production and ethics, Mickel and colleagues (2023) conducted mixed-methods lab ethnographies to explore what they call lab-specific 'cultures of ethics' (p. 4) or 'ethical cultures' (p. 5). They define them as the lab-specific ordering of preferences and issues that determine how ethicality is enacted as part of the researchers' daily practices (Mickel et al., 2023). They spent prolonged time with and investigated the ethical cultures in a bioengineering laboratory and two psychology laboratories. Due to the comparability to my ethnography, I will here focus on their work in the psychology groups. In the psychology labs there were two fundamental themes underlying the ethical considerations of the research. First the motivation to obtain clean and good data from participants. Hence, good experimental design is seen as the proper combination between epistemological and ethical considerations (Mickel et al., 2023). One of the aspects that came up as influencing the quality of the data is how to instruct the participants to ensure that they act as required and are motivated to aid the researchers in scientific knowledge production. Second, ethical considerations of their research seem to be taught and instructed during lab meetings. According to Mickel and coauthors (2023), this conviction that ethical research can be taught and instructed was indicated by frequent reference to rules and institutional frameworks and stems from an institutional context with strict and salient ethical guidelines and standards. Hence, STS studies suggest that there is a co-production between ethical and epistemological considerations in the process of scientific knowledge production within the specific context (e.g., country, discipline, institute, lab) in which the practice of scientific research is enacted (Gläser et al., 2015; Jasanoff, 1996; Mickel et al., 2023; Pickersgill, 2012).

As the last chapter already demonstrated, when thinking about what research integrity might mean for the researchers, the topic of research ethics comes up. Furthermore, in the neuroimaging lab in Germany and in the biopsychology lab in Germany I was sometimes able to observe the researchers' interactions with the participants. These interactions were characterised by certain reoccurring social interactions. For example, while participants were in the MRI (magnetic resonance imaging) machine, participated in TMS (transcranial magnetic stimulation) studies, or pupillometry and EEG experiments (electroencephalogram) the researchers were quite attentive towards how the participants were doing and feeling. This was demonstrated by repeatedly asking things like, how the participants are doing, whether everything is okay, whether they would like to take a break, and whether they would like a snack or something to drink. Furthermore, there was usually some small talk between participants and researchers, especially during the phases where participants were prepared for the measurements with the instruments. Therefore, during the interviews I asked the researchers, what the role of the participants and their relationship with them is in their research and the ability to achieve 'good research'. Moreover, since my field observations took place during social interactions between scientists in situations where they talk about their research, and the lab climate was brought up by some scholars spontaneously, the type of social relationship the researchers in the groups have to each other and how that influences their experience and research (De Vries et al., 2006), became another point for conversation during the interviews. An enactment of research as a co-production between epistemological and ethical considerations also became salient during my interviews (Mickel et al., 2023; Pickersgill, 2012). The following quote conceptualises research integrity as the combination of doing good science and being ethical while doing it:

> There are sort of two parts to it. One is you're doing good science. And the other is … You're not hurting anyone while you're doing your good science. And that includes both members of your lab, but also experimental participants (Reader Cognitive Science, The UK).

On the one hand, doing 'good science' is described as 'making a good faith effort to reach the truth and to help your wider scientific community reach the truth' (Reader Cognitive Science, The UK) by not fabricating data, using wrong statistical analysis or making unsubstantiated claims in the context of the formal assumptions. On the other hand, being ethical is conceptualised as treating members of the lab and participants well. According to the researcher, this specifically contains 'not playing stupid authorship games', 'not being a bully to people in your group' (Reader Cognitive Science, The UK), not inappropriately taking credit, not doing harm, and avoiding deception. Therefore,

separating 'good research' and research ethics can become problematic when attempting to understand research integrity as 'good research' from within, especially in the context of research involving human participants which is described as a co-production of ethical and epistemological considerations (Mickel et al., 2023; Pickersgill, 2012). Therefore, in this chapter I examine two key aspects of ethical scientific conduct, to investigate their relationship with research integrity. One is the ethics in the context of how to treat participants and the social rapport with them while the other is the treatment of and relationship with one's colleagues in the lab.

The ethics of dealing with participants

Similar to Mickel and colleagues' (2023) findings in the psychology lab 'the ethical culture was structured around dual principles of caring for science and for human subjects, where producing reliable, "clean" data was not only an epistemological concern but also an ethical' one (Mickel et al., 2023, p. 4). Hence, the ethicality of their research conduct with participants seems motivated by two poles, the well-being of the participants and the quality of the data. This can also be conceptualised as ethicality in research involving human participants having a deontological and a teleological component (Israel, 2015). According to teleological or consequentialist ethics, the conduct is judged based on the specific outcome it is supposed to achieve. Thus, what is useful for achieving a specific goal is seen as good or right. In deontological ethics, the goodness or rightness of behaviour is judged based on its alignment with specific principles or rules (Israel, 2015). Whether those are on equal standing or the comfort of the participants is the proximate goal serving the ultimate goal of receiving high-quality data remains an open question. During my observations and in conversations with the researchers I noticed that on the one hand there is the teleological part where building rapport and a good relationship with the participants is seen as a skill that facilitates getting good, high quality and reliable data by ensuring that the participants are motivated and properly follow the instructions of the scientists during the studies.

> When I'm collecting data, I take great care of my participants. In the end, the data comes from them. So, you want them motivated. You want them to feel respected. I try to build a quick relationship with the participants […]. I feel really responsible, also for their comfort and that they are at all times, not only engaged, because then I get good data quality, but also that they are liking what they are doing. […] I do notice when I have a good relationship with the participants, they will work harder for me (PhD Neuroimaging Lab 1, The Netherlands).

In the preceding quote the researcher also addresses how, according to their own experience, the way they treat or engage participants can help in addressing the inherent sensitivity and variability of the measurement and data collection process in imaging studies. The PI of Neuroimaging lab 1 in the Netherlands supports that assertion and adds that this social aspect of the research involving human participants is sometimes underappreciated while also explicitly mentioning the crucial importance of building rapport with the subject to ensure proper and motivated participation and ultimately good data.

> I think that's sometimes a bit of an underappreciated aspect of data collection. Where people think, well, if I coded everything and I thought of my paradigm well, then everything is fine. [...] The way I see participants is really as a, almost as a collaborator in a way to the experiment. [...] If their motivation is going to vary a lot over the experiment, that's going to introduce a lot of variation in the data that is again, not related to the conditions that we manipulate. [...] And in order to minimise the effect of those, I think it is really important that participants realise that they should try to maintain a high and constant state of alertness. And those things you get from having a good rapport with the subject (PI Neuroimaging Lab 1, The Netherlands).

Subsequently that PI further introduces a reflexive moment, by comparing the relationship and rapport between them and myself during my fieldwork in their labs, and especially the interview situation, to the type of relationship they build with their participants in their own research to ensure a willingness to work towards a clean and motivated data collection process:

> Just like we now sort of have an implicit bond where I try to help you and think along with you. But if we would not have that, then you would not get the data from me that you are hoping for (PI Neuroimaging Lab 1, The Netherlands).

In the German neuroimaging lab, I was also able to observe an interaction between a postdoc and the lab manager, as the excerpt from my fieldnotes shows.

> The lab manager had recruited a participant on the phone and told them that the specific study was about investigating a specific cognitive capacity. After which the postdoc told the lab manager, that it should have been phrased differently, and they are supposed to tell the participants that the study is about improving this cognitive capacity instead of just investigating it.

When I asked the postdoc about why this distinction is so important when talking to participants, the postdoc replied that 'it helps when it sounds more interesting, because we are then rather able to get them excited, than with a standard address' (Postdoc Neuroimaging Lab, Germany). This is also described as important in the recruitment process because there appears to be a shortage of available participants for neuroimaging studies (van Drimmelen

et al., 2025 describe such a shortage of participants in end-of-life research as one source of uncertainty in researchers' discretion).

On the other hand, as some of the quotes already indicated, there is the deontological component where researchers describe the act of making the experience of participating in their research as comfortable and positive as possible for the participants as their duty as well as responsibility as scientists.

> I also believe that if it were completely irrelevant [...] for the research, it is also important for the test subjects or for the social relationship between the test subjects and the researchers. [...]. So it's all the more important to create a dignified, respectful atmosphere. Even if that might be completely irrelevant for the research results. [...] But nevertheless, I suspect that the quality of the data rather benefits from this, perhaps not necessarily[36] (Professor Biopsychology Lab, Germany).

This researcher emphasises that even if the method of engaging with and treating participants does not affect the quality of the research or the data collected, it remains their responsibility to ensure their participation is as 'dignified' and 'respectful' as possible. However, it is important to keep in mind that the presented deontological component might just be an effect of social desirability and wanting to preserve a certain image of their practice to the outside (Hangel and Schickore, 2017; Mulkay, 1976; Tavory, 2020), while the primary motivation and function behind the kind of treatment of and rapport with the participants could just be their utility in achieving the researchers' goal of obtaining clean and useful data.

Most of the utterances from the researchers addressing the social relationship and rapport they build with the participants directly illustrate an intricate combination of deontological and teleological considerations concerning the ethics of their research regarding the experiences and state of mind of the participants, as if they are co-constitutive of each other.

---

[36] Original in German: ‚Das ist auch glaube ich, also wenn das völlig irrelevant wäre, wie wichtig es für die Forschung ist, ist es auch wichtig für die Probanden oder auch für das Sozialverhältnis zwischen Probanden und den Forschenden. Und die Rolle ändert sich ja ständig sozusagen. Also wir alle sind ja sowohl Probanden als auch Forschende. Die Studenten sind am Anfang vorwiegend Probanden, aber später machen sie auch eine [Forschungsarbeit]. Die Freunde der Studenten sind ja nur die Versuchskaninchen. Also damit ist es umso wichtiger, dass man eben sozusagen, dass es einfach eine würdevolle, respektvolle Grundatmosphäre ist. Auch wenn das vielleicht für das Forschungsergebnis völlig irrelevant sein könnte. Auch dann muss man das trotzdem genauso machen, wie wir es machen. Oder vielleicht noch besser, vielleicht machen wir das gar nicht gut genug. Aber trotzdem ist es natürlich auch, ich vermute schon, dass die Datenqualität da eher profitiert davon, vielleicht nicht notwendigerweise.'

> we want them to feel comfortable. They serve us. First of all, because especially in the EEG area, if test subjects are tired or demotivated, the data quality is poor. I have an intrinsic interest that they feel comfortable and, I also consider it my duty to make them feel comfortable when they're here in our lab. So that they are in safe hands. [...] It's essential to me that people are also safe in every situation. So, both from a technical and a hygienic or other point of view[37] (Lab Leader Biopsychology Lab, Germany).

As illustrated by the last quote the importance for contemplating how to engage with and treat participants is often presented in the context of expressing a sense of indebtedness to and dependence on the participants as researchers, because without their voluntary participation in their studies, they would not be able to actually do their research at all: 'Without participants, we couldn't do this' (PhD Neuroimaging Lab 2, The Netherlands). In other words, the feeling seems to be that they owe it to the participants to treat them well and ensure their comfort as well as well-being: 'that's why it's important to be mindful of the participant experience' (PhD Neuroimaging Lab 2, The Netherlands). Therefore, the actual execution of their research and the collection of data is presented as a cooperation between them and the participants. They seem to be in it together and rather than mere data sources participants are seen as collaborators in the process of data production. As a previous statement suggests: "I see participants really almost as a collaborator' and another scholar succinctly put it, when asked about the role of the participants in their research: 'I think the responsibility is shared' (PhD Neuroimaging Lab 2, The Netherlands)

---

[37] Original in German: ‚Ja, weil wir wollen, dass sich die wohlfühlen. Die dienen ja uns. Also erstens, weil es insbesondere im EEG-Bereich, wenn Probanden müde oder demotiviert sind, ist die Datenqualität schlecht. Also von daher habe ich ein intrinsisches Interesse, dass sich die wohlfühlen und, äh, ich halte es auch für meine Pflicht, wenn sie hier in unserem Labor sind, dass sich die wohlfühlen. Also dass sie in sicheren Händen sind. Also auch, wie wir, also auch am anderen Ende des Spektrums, in dem wir, also ich bin der Laborleiter des Labors und, ähm, äh, es ist mir essentiell wichtig, ähm, dass die Leute auch sicher sind in jeder Situation. Also, sowohl, von technischer als auch von hygienischer oder sonstiger Seite. Wir machen keine klinische Forschung, die Leute haben keinerlei Nutzen. Von dieser Forschung, sondern das ist, also ich bin akademischer, Akademiker und ich freue mich an dieser Freiheit zu forschen, an dem, was mich interessiert. Ich habe aber so ein gewisses Grundverständnis, dass andere Leute nicht verstehen können, dass das interessant ist, was ich erforsche. Und, äh, von daher, äh, muss ich dafür sorgen, dass es den Leuten gut geht. Weil die sind da sozusagen halt verpflichtet, zum Teil eben, weil viele Versuchsstunden bekommen. Also, Kurscredits und, äh, das ist ja dann nur bedingt freiwillig, also ich meine natürlich, dass nach im legalen Sinne freiwillig, die können da woanders hingehen. Aber sie müssen irgendwo ihre Versuchsstunden ableisten. Und bei uns kriegen sie halt viele auf einmal, weil unsere EEG-Experimente relativ lange dauern. Ich glaube, drei Stück und, äh, darum kommen die halt gerne zu uns. Aber wir wollen ihnen unter gar keinen Umständen schaden.‘

> We are super dependent on the participants and it's also super important that they take it seriously what we do. And if we don't get reliable data, we can't do anything. […] And we really try to also stick to all those ethical guidelines and have there also quite strict rules how to do all of this (PhD Neuroimaging Lab 2, The Netherlands).

Furthermore, similar to the culture of ethics described in the psychology lab(s) by Mickel and co-authors (2023) in these neuroimaging labs there seems to be a salient and strict influence exerted by institutional ethical standards and ethical frameworks, like requirements set by institutional review boards and the project presentation meetings, where prior to being allowed to use the machines one has to present their study plan and justify the usage of machines to members of the institute.

According to the researchers, the participation in their research can be quite taxing for people and any kind of deviation from the instructed behaviour or movements in the scanner can introduce unwanted noise or variation into the measurement process and the data. This apparently important aspect of how to instruct participants also came up numerous times during the field observations, especially during supervision meetings between the PI's and the PhDs as well as postdocs of the neuroimaging labs. Next to written or protocolised and somewhat standardised instructions, the way participants are treated and handled during the actual study situations is seen as a crucial aspect of their research interaction with participants that at least partially enables the researchers to control and alleviate these disturbances resulting from the nature of their research, the data collection process, and the dependency on the cooperation from the participants. Therefore, this skill of building rapport is also seen as one potential way to address the fact that the participants in combination with the sensitivity of the measurement instruments are a critical source of uncontrolled variability. It is explicitly mentioned that participants are not standardised units, but they differ. Therefore, there is no standard way of engaging with them. Hence to a certain degree it is a result of social interaction and tacit skill instead of following research protocols:

> No, not like I have a standard approach […] On the go, I have to sort of assess what kind of person is this and how can I best make them comfortable? […] you have to be a bit sensitive to this (PhD Neuroimaging Lab 1, The Netherlands).

The aspect of building rapport with the participants in the form of a good and comfortable social relationship is usually not heavily discussed or focused on in method textbooks and other training material for quantitative research. This skill or component of the research methods is however extensively discussed in the context of qualitative research (see e.g., Duncombe & Jessop, 2002; Willig & Stainton-Rogers, 2017). There it is described as a skill and

practice that intends to ensure that participants stay engaged during the research situation and provide honest as well as sincere answers to questions, for instance in interviews (see e.g., Duncombe & Jessop, 2002; Willig & Stainton-Rogers, 2017). Hence, in qualitative research the link between the ethical relationship and good rapport that the researchers have with the participants to the quality of the data and ultimately the research seems more salient or explicit and is an established consideration for good qualitative research practices. In contrast, a recurring theme during my investigation that emerges from the researchers is that how to properly treat participants and establishment of good rapport with them, while being seen as a necessary skill and crucial component of their work, researchers were not aware of it prior to having conducted this type of research themselves, or it is sometimes even hidden.

> I think it's maybe something that people sometimes try to hide even a little bit, because they want to kind of make it seem as if, as long as your experiment is good, then the participant is just kind of like a necessary element of it. That it's like a machine and it will always do the same thing. But that's just really not how humans are (PI Neuroimaging Lab 1, The Netherlands).

This suspicion of attempting to hide this social and subjective dynamic inherent to their research could be seen as aiming to preserve an image of objectivity as a medium for the communication of credibility within the research community (Daston, 1995; Daston & Galison, 1992, 2007; Field & Derksen, 2021; Penders, 2022). One could argue that in the community what is seen as credible in reporting and communicating of research as objective and ordered is obfuscating what is relevant and valued as 'good research' practice in the laboratories (Lynch, 1985). Hence, building rapport and social interaction with participant can be seen as an invisible component of good research practice (Pols et al., 2024).

Probably because the skill of treating participants well and establishing rapport with them is typically not a salient or large part of the quantitative methods curriculum nor of the reporting of the research, it is described as a tacit skill that remains as personal knowledge of the researchers or in the confines of the specific lab (culture) (see e.g., Mickel et al., 2023). This intricate tacit knowledge is acquired through the experiences and socialisation as a researcher in a specific lab environment and learned based on the repeated interactions with participants (Collins, 1974; 1985; 2010; Delamont & Atkinson, 2001; Polanyi, 1958). A PhD student from the German neuroimaging lab further describes this as 'one of those ancillary skills that you would never think about' and put a caveat or even a limit on how comfortable participants should feel during the research to ensure participants' 'good experience and a good experiment' (PhD neuroimaging lab, Germany). According to this researcher

it is a 'constant tweaking' (PhD neuroimaging lab, Germany) to find the required balance, regarding how to engage with the participants, between being authoritative as a scientist and being friendly to ensure their comfort. Hence, the ethical treatment of participants is not presented as a standardised practice or the result of clear-cut rule following, but it is rather described as a situational enactment of ethics in practice that is grounded in the socialisation of researchers into the routines and ways of the lab and requires tinkering in reaction to the idiosyncrasies of the specific participant-interactions (Knorr, 1977; Knorr-Cetina 1981; Mickel et al., 2023).

The ethical colleague

While the role that the institutional culture or climate plays in affecting research integrity has recently received more attention (see e.g., de Peuter & Conix, 2023; Forsberg et al., 2018; Labib et al., 2024), a critical investigation on a finer level of granularity, like individual labs, is still largely missing. Usually, the focus is on the organisational climate, for example, as defined by Schneider and co-authors (2013), and how it can be conducive or detrimental to research integrity (see e.g., Armond & Kakuk, 2022). Furthermore, Haven and co-authors (2020) conducted focus groups with researchers to inquire into their views on characteristics of a responsible research climate. They identified six core characteristics: fair evaluation, openness, sufficient time, integrity, trust, and freedom. However, it is questionable whether those six are really distinct characteristics and all of them remain quite elusive. Moreover, every time the term integrity itself turns up as a defining characteristic for anything to do with research integrity or responsible conduct of research it borders on tautology, because responsible conduct of research and research integrity are so conflated and elusive themselves in the discourse (Field et al., 2024; Kolstoe & Pugh, 2024). Additionally, some exceptions do exist, but they do primarily focus on the role of leadership in the lab, or supervision (see e.g., Haven et al., 2020, 2023; Resnik et al., 2023, 2024). Research integrity is identified as one of the motivating factors for managing the ethics of the lab (Resnik et al., 2023). Furthermore, Resnik and colleagues (2024) discuss ethical decisions under the lens of role ethics, due to the different roles scientists can take on as researchers, mentors, entrepreneurs, and managers. Accordingly, due to the various differing obligations inherent to those roles, ethical dilemmas can emerge affecting their daily practice as a scientist (Resnik et al., 2024; also see van Drimmelen and colleagues, 2025 for similar elaboration concerning moral ambivalence due to conflicting parallel existing values of proper conduct).

Teamwork and a supportive working relationship with lab colleagues are established characteristics of the research climate to consider in discussions

related to issues of research integrity (see e.g., Blatch-Jones et al., 2024; Brenninkmeijer & Pols, 2026; De Vries et al., 2006; Pols et al., 2024). Therefore, here I will explore how the researchers relate experiences of the lab climate to issues surrounding research integrity as 'good research'. Here the lab climate is meant to refer to how comfortable or pressured researchers feel within the direct environment of their research group. A recurring theme is the crucial importance of ensuring a research climate that makes everyone in the lab group feel comfortable enough to speak out and admit errors or mistakes as well as voice criticism towards colleagues (see e.g., Frank et al., 2026). Within the lab context, the research climate can function as somewhat of a catalyst or even precondition, but also as a barrier or inhibitor for the aforementioned component of research integrity of being critical and reflective (reflexive) (Field et al., 2024; Valkenburg et al., 2021). In the following quote a postdoc goes exactly into this and elaborates on their responsibility for ensuring that the scholars in the lab feel comfortable and know that they can come and talk about something that might not have worked or when they might have any worries about one's trajectory in academia. This aspect of worries about one's trajectory in academia becomes especially relevant when considering that, especially in the Netherlands, the most common answer to the question, what they would change in their direct environment that would allow them to conduct their research better, was to abolish the short-term contracts of postdocs (see Appendix A). Similar to earlier comments from a Dutch postdoc, conduct that is perceived as deviating from norms and values of scientific practice are depicted as ways to navigate uncertainty and stress in science (see e.g., Elizondo & Kaltenbrunner, 2024). While the lab climate is presented as a factor that can contribute to researchers not choosing this way out of such situations.

> It is very important to create an atmosphere where people actually feel comfortable to speak out. Where they know that when they have a bug in their code, and all their results actually are very different from what they thought, that they will not try to hide that from me, but will be able to say, oh I messed this up and its fine and we just do it again. That they know that they can trust me and that if they have any worries about their science, about their career that they can just come to me and talk about that. Things will not work out in science you will find different results, your hypothesis will not come true, and this can be really frustrating, and you can perceive yourself as a failure. And think to never have a chance in science, personally that can be very stressful and that can then of course also impact the quality of your research, where you might then consider to just produce some fake results (Postdoc Neuroimaging Lab 1, The Netherlands).

Among similar lines, in an attempt to explain the role a lab climate plays, where 'you can trust others, both colleagues and superiors, to the extent that

you have the confidence to address everything, i.e. doubts and also when things aren't going well' (PI Neuroimaging Lab, Germany), the PI of the German neuroimaging lab, recounts a 'toxic' lab climate that the PI and a postdoc experienced in a previous lab and how that lead to a climate of 'fear' where people stop 'daring' to speak up or 'say everything'.

> So, that you somehow didn't like going to these meetings, because it was constantly escalating [...] of course that leads to you no longer daring to say everything you think and, above all, no longer daring to introduce things. [...] [A student] said of their own accord, I want to present here and I think that's totally great. And, of course, that's not possible if there's such an atmosphere of fear or, as you also hear from some labs, if something goes against the theory that the boss advocates, if it doesn't agree with it, then it's often swept under the rug or not allowed to be published[38] (PI Neuroimaging Lab, Germany).

This social component of the everyday research reality in the context of one's lab seems so important that it is also described by some as a crucial consideration in hiring situations to ask whether the personality of the applicant

[38] Original in German: ‚Also ich würde sagen, vertrauen, dass man den anderen, also sowohl den Kollegen als auch den Vorgesetzten so vertrauen kann, dass man sich traut, alles anzusprechen, also Zweifel und auch, wenn es nicht gut läuft. Und [postdoc] und ich haben ja gemeinsam in der Neuropsychologie angefangen und da war eine Zeit lang nicht so gute Stimmung, weil es ein paar Mitarbeiter gab, die nicht miteinander harmoniert haben. Die waren persönlich ein bisschen schwierig und das waren halt einige und dadurch herrschte [postdoc] hat mir das danach erst gesagt, dass [postdoc] das immer so als toxische Atmosphäre empfunden hat und das fand ich eigentlich ganz treffend. Mir ist das dabei, also ich fand es immer eigentlich nervig, aber erst so im Nachhinein, als [postdoc] das so gesagt hat, dass [postdoc] die Meetings immer schrecklich fand, die Gruppenmeetings, weil das so toxisch war, ist mir das aufgefallen, dass das, also das ist ja ein sehr starker Begriff, aber ich glaube, da hat es tatsächlich, da traf das wirklich zu, sodass man irgendwie ungern in diese Meetings gegangen ist, weil es ständig eskaliert, ist und man schon immer versucht, ich bin dann da öfter eingeschritten, weil ich da auch schon Gruppenleiterin war und habe versucht, dann ja, da so ein bisschen zu schlichten und die eine Person, die damals auch noch Doktorand/ Doktorandin war, ein bisschen zu schützen, weil die mir so leid tat, also die da wurde dann auch schon mal geweint und sowas und das führt natürlich dazu, dass du dich nicht mehr traust, alles zu sagen, was du denkst und vor allem auch nicht mehr traust, Dinge vorzustellen. Und das finde ich bei uns ganz gut, zum Beispiel gestern, im Lab-Meeting hat ja auch [jemand] vorgestellt, [eine Person die] gerade erst sein Masterstudium abgeschlossen hat und für die, wenn sich [postdoc] da hinstellt, ich glaube, [die Person] ist nicht groß aufgeregt, aber für jemanden, der/die noch so am Anfang von der Karriere ist, ist das total die große Sache und dass die trotzdem von, also [diese Person] hat von sich aus gesagt, ich möchte hier vorstellen und das finde ich total super und das geht natürlich nicht, wenn so eine Atmosphäre der Angst herrscht oder, was man auch natürlich aus einigen Labs hört, wenn etwas gegen die Theorie, die der Chef vertritt, wenn es da nicht mit übereinstimmt, dann wird das dann auch gerne mal unter den Tisch gekehrt oder darf nicht publiziert werden. Und also das ist ja ganz schlimm.'

actually fits into the current lab group. The next quote further suggests a link between such a comfortable lab climate and being able to do 'good research':

> It is a big component that you feel comfortable within your closest surroundings to be able to relate to them, to be able to contact them for questions, for concerns when you don't feel you're doing well. […] I think that without that, I don't think it allows you to do good research if you're just isolated […] And so I think it is key, both for social support, but also for actual work support for actual scientific questions (PhD Neuroimaging Lab, The Netherlands).

The researchers also voice the sentiment that being a researcher and conducting research can be quite the insulated and 'lonely activity' (Postdoc Cognitive Science Lab, UK), but having a lab climate that is conducive to constructive critical feedback is seen as enabling a valuable form of social interaction within the lab that fosters not just 'good research' but also proper mental health (Troughton & Obasi, 2022; see Lubega et al., 2023 for issues of PhDs' mental health in the context irreproducibility). In that context a scholar talks about the influence of a good lab climate in terms of providing a sense of safety and sharing a responsibility to take care of each other.

> A good research climate with your lab is determining your psychological well-being. […] If you feel that you can share your research and you know you will get critical comments, but it's also going to be constructive, then you're also more open to this type of feedback and that will improve your research. In that sense I think research is a very weird endeavour, because it's all these little islands, everybody's doing their own projects, but you're still sharing it with each other. […] We also have a common responsibility of taking care of each other not just our well-being but also our work. It's really important actually that you are able to reflect on your work with your group and they can help you reach a higher level. And that requires sort of a safe environment (PhD neuroimaging Lab, The Netherlands).

One component of this openness to criticism as part of a good lab climate seems to be the absence of shaming people when something does not work out as intended or for making mistakes and errors (Frank et al., 2026). Hence, this seems to hint towards Macleod's call to stop playing the 'blame game', where the focus is less on punitive action and holding individual researchers accountable, but more on the collective and the institutions to facilitate better research instead of (only) punishing bad conduct (Macleod, 2021, p. 533).

> If you're in an environment where it's not okay to make mistakes or when people think, don't talk about it, so there's this implicit sort of shame. Then it would, I think, have big consequences for the research that you're doing (Postdoc Neuroimaging Lab, The Netherlands).

However, while most researchers seem to believe that a good lab climate and a proper relationship with the colleagues is conducive to research integrity, some also complexify the issue by pointing towards something of a dilemma between the two. They, for instance, suggest that by strictly demanding practices that can be seen as promoting research integrity the climate in the group could actually be harmed. While prioritising the climate in the group over proper adherence to standards and rules for proper conduct can counteract research integrity. The following researcher provides the example in the context of authorship and calls for treating these issues as a balancing process, because one has to decide at some point which of the two are deemed more important. While practically equating Open Science with research integrity, this scholar even suggests that research climate and research integrity might be independent of each other.

> If I say [to the doctoral candidate], well, that wasn't enough for a first authorship and withdraw it, then that is an improvement in the Open Science perspective or integrity. But it will certainly worsen the research climate. […] I do believe that enforcing research integrity, so to speak, can have a negative impact on the research climate. And that you have to consider what is more important to you. It is always a balancing process. […] I think I can imagine Open Science even in working groups where there are perhaps individual groups, where there is not a good research climate, I can imagine that they do very good Open Science. Or I can also imagine in them, oh, we are fantastic, we do great, we all know each other well, we are all great, we all take each other on as co-authors, of course. Yes, that's an example of a good research climate, I have heard that too. I have heard that from the working group. Great research climate, from a research integrity perspective, I am trying to use the term now, it's quite bad. […] I would say independent[39] (Professor Personality Psychology, Germany).

[39] Original in German: ‚Es ist halt schon schön, wenn so ein Arbeitsklima gut ist, würde ich jetzt mal denken. Aber es gibt natürlich alle möglichen Gründe, warum das vielleicht auch mal nicht so ist. Aber ich weiß ja, vieles ist erstmal unabhängig. Von jetzt an Qualität der Forschung, wie man das durchführen kann. Wenn ich zum Beispiel [der Doktorantdn/ dem Doktoranden] sage, naja, das war jetzt aber nicht genug für eine Erstautorenschaft und das entziehe, dann ist das eine Verbesserung der Open Science Perspektive oder der Integrität. Wird aber sicher zu einer Verschlechterung des Forschungsklimas. Würde ich jetzt sagen. Insofern sind vielleicht auch manche Sachen, die man macht, um den Forschungsklima noch okay zu halten. Ja, das würde auch bei so Autorenschaft. Ich hatte das auch schon mal. Da hatten wir einen Autor, da war einer dabei, der hat immer ein bisschen was geschrieben, ein bisschen was überarbeitet. Ich fand es eigentlich dann nicht genug für eine Autorenschaft. Dann gab es ein Nachfolgeprojekt. Eigentlich war der da schon dabei, weil das halt am Anfang an ja auch so ein bisschen deren Projekt war. Und dann habe ich halt irgendwie so eine E-Mail geschrieben und so gesagt, ich glaube, das ist jetzt okay, wenn wir es jetzt beim nächsten Mal oder beim nächsten Mal nicht, dabei ist. Es war jetzt auch nicht genau das Projekt.

As the previous quote already suggests this potential conflict between a good lab climate and research integrity came up during my fieldwork especially in matters of authorship. A postdoc in a Dutch neuroimaging lab further elaborates on the somewhat ambivalent relationship between authorship and the lab climate. The scientist explains how it can sometimes be 'tricky' to deal with situations where it is hard to distinguish whether someone is just helping out a friend or officially a co-author on a paper. Moreover, the interlocutor highlights the social nature of these research situations (De Vries et al., 2006), by suggesting that dealing with these ambivalent situations is 'a matter of social skills'.

> Something that can be tricky is that actually, paradoxically, especially if the lab environment is good and you really are friends with your colleagues, is that people can ask for favours: Could you help me with this? […] the favours pile up and at some point, if you cannot say no to a friend, it can get weird or annoying. […] But it's a matter of social skills, right? Because also you want to help your friends. you have to keep an eye on how much time you're investing in someone else's project (Postdoc Neuroimaging Lab 1, The Netherlands).

Also ich habe so ein bisschen versucht, noch nett zu umschreiben oder sowas. Im Nachhinein glaube ich, ein Fehler. Ich würde es auch nicht mehr so machen. Ich würde es jetzt immer irgendwie versuchen, irgendwie telefonisch zu klären oder irgendwie zu klären. Wahrscheinlich würde ich dann wieder weiter nächstes Mal mitschleppen. War wahrscheinlich mit dem Forschungsklima zu der betreffenden Person, glaube ich, nicht so gut. Also ich glaube schon, dass sozusagen Durchsetzen von so Forschungsintegrität aufs Forschungsklima sich ausschlagen negativ auswirken kann. Und dass man halt überlegen muss, was ist einem wichtiger. Das ist immer ein Abwägungsprozess. Andersrum, Forschungsklima zu Forschung, weiß ich jetzt nicht, das andersrum wirkt, naja, wenn das Forschungsklima halt irgendwie, ja klar, wenn natürlich, wenn das Forschungsklima jetzt wie sowas ist, wenn der Vorgesetzte sagt, publizieren, publizieren, publizieren, ist mir doch gerade egal, ob du Open Science machst oder nicht. Aber das ist ja nicht, dann ist es natürlich ganz schädigend, wenn der Druck ausübt und sagt, aber der Druck muss reproduzierbar sein und das muss so und so und das muss so und so und klare Vorgaben gibt und aber dabei sehr direktiv ist, ist es halt auch schädigend für das Forschungsklima. Aber gut für Open Science. Also ich sehe das, vielleicht sehe ich das ein bisschen ähnlich wie die Diversität. Ich kann mir glaube ich Open Science auch in Arbeitsgruppen, wo es vielleicht einzelne Gruppen gibt, kein gutes Forschungsklima existiert, ich kann mir vorstellen, dass sie sehr gute Open Science machen oder, ich kann mir auch darin vorstellen, ach, wir sind fantastisch, wir machen super, wir kennen uns alle gut, wir sind alle super, wir nehmen uns alle miteinander drauf als Co-Autor, selbstverständlich. Ja, das ist ja ein Beispiel von gutem Forschungsklima, habe ich auch schon gehört. Ich habe das schon gehört von der Arbeitsgruppe. Super Forschungsklima aus einer Forschungsintegrizitätssicht, ich versuche jetzt das mit dem Begriff zu nehmen, ist ja ganz gut, ist es schlecht. Autorenschaft nur wegen, weil man einen Teil eines, eine Arbeitsgruppe mit gutem Forschungsklima ist, ist glaube ich nicht ausreichend. Also insofern würde ich sagen, unabhängig.'

Relatedly, one researcher introduces a distinction between working climate and research climate. With research climate including things like 'friendliness for research, academic freedom', and 'restrictions' in form of, for instance, 'time pressure'. In contrast working climate, refers to the fact that like any other occupation being a scientist is a job. Therefore, the relationship to the colleagues does not necessarily have to reach the level of a personal relationship in the form of something like friendship, but it has to be professional.

> Working climate is one thing and research climate is something else for me. Research climate is simply the friendliness for research, for example towards teaching. As one aspect and research as freedom of research/academic freedom is something that is part of the research climate for me. So, can I research what I want, or do I have to research what the boss wants? And what are the restrictions, in terms of integrity, time pressure, whatever, that's research climate for me. And working climate is something completely different for me. Also, this is my job, and, I don't have to hang out with my work colleagues in the evenings, if I don't want to, that's fine, if that's the case. But we're in a professional environment, just like in a company. I want to deal with people professionally, and, I don't have any demands that I have to be friends with them, with everyone[40] (Lab Leader Biopsychology Lab, Germany).

While the aforementioned perceived 'publish or perish' pressure in academia also arises in discussions about lab climate, and is generally seen as negatively correlated with research integrity due to related issues such as workload and time pressure, researchers are also aware that there is a limit to how much the lab climate can alleviate this pressure, given the prevailing pressure at the systems level and the researchers' interdependence on several levels with different stakeholders within the transepistemic arena (Latour & Woolgar, 1979; Knorr-Cetina, 1982; Smith, 2014; Whitley, 1984):

---

[40] Original in German: ‚Ich würde jetzt zwei, ähm, ähm, ich höre da was raus, ich weiß nicht, ob das da drinsteckt, äh, oder nicht. äh, was ich da raus höre ist, also erstens, Arbeitsklima ist was eines und Forschungsklima ist für mich was anderes. Und, ähm, Forschungsklima ist einfach die, die, die Freundlichkeit für Forschung, zum Beispiel gegenüber der Lehre. Als einen Aspekt und Forschung als, ähm, eben Forschungsfreiheit, ist was, was für mich zum Forschungsklima gehört, also, darf ich forschen, was ich will, oder muss ich forschen, was der Chef will? Äh, und, äh, was sind da die Resektionen, eben was Integrität, Zeitdruck, sonst was anbelangt, das ist Forschungsklima für mich, und Arbeitsklima ist was für mich was völlig anderes. Und da habe ich auchalso, ich, ich will, ich, das ist mein Job, und, äh, äh, ich muss nicht mit meinen Arbeitskollegen abends, abends abhängen. Wenn ich das nicht möchte, das ist schön, wenn das ist. Aber wir sind in einem professionellen Umfeld, äh, genauso wie in der Firma auch, und, ähm, ich möchte professionell mit den Leuten umgehen, und, äh, ich habe keine Anspüche, dass ich mit denen, mit allen, äh, Freund sein muss, sondern ich muss professionell und, urteilsfrei umgehen, aber, keine persönlichen Beziehungen deswegen notwendigerweise aufbauen.'

> But we are very strongly led to believe that our trajectory depends on publication. So, it's nice if the PI doesn't pressure you, but everyone is still pressured from the outside. So, it's hard (Postdoc neuroimaging Lab 1, The Netherlands).

Discussion and conclusion

The ethicality of research conduct involving human participants, is described as being driven by a certain 'duality of caring' (Mickel et al., 2023, p. 4). On the one hand researchers suggest that they view it as their duty and ethical obligation to ensure the participants' well-being and comfort. On the one hand, this consideration seems to be fuelled by, for instance, a sense of indebtedness to the participants, because without their voluntary participation and cooperation there would be no such research and the adherence to specific institutional ethical standards regulating and guiding research conduct as well as the treatment of participants in combination with a sense of duty or responsibility for their well-being. On the other hand, the interlocutors indicate that building rapport with the participants and caring for them is serving the ultimate aim of getting high quality data, since the data comes from the participants (Mickel et al., 2023; Peterson, 2016). Hence, due to the combination of the sensitivity of the measurement instruments, the demanding conditions of participating in imaging studies as well as other neurosciences, and the high variability introduced into the research by the human subjects, how to engage with the participants becomes a tacit skill that at least partially addresses these issues linking ethical considerations through their influence on data quality to research quality and ultimately research integrity considerations (Helgesson & Bülow, 2023; also see Peterson, 2016 for similar observations in developmental research). However, as the respondents indicate, this aspect of good research remains often hidden or invisible (Pols et al., 2024), presumably behind an image of objectivity and order to convey credibility to the community (Daston & Galison, 1992, 2007).

The other component of the ethicality inherent to the daily practices of the researchers is the type of relationships and social interactions they have with colleagues, especially in the direct lab environment – here called lab climate. A good lab climate is described as ensuring the possibility for critical and constructive exchange that enhances the research and the personal well-being of the researchers (Field et al., 2024; Valkenburg et al., 2021). Considering that researchers talked in Chapter 5 about the importance of being able to critically reflect on research individually and in their lab group, this links these ethical considerations to researchers' understanding of research integrity understood as 'good science'. Moreover, a good social and comfortable lab climate is seen as enabling the sharing of responsibility for the research that comes out of the

lab, instead of being on the shoulders of just one person, supporting mental health, and providing a sense of safety that can lead scholars to seek advice or help, instead of misbehave in situations of uncertainty (Elizondo & Kaltenbrunner, 2024). However, at the same time the lab is situated within the larger context of the research ecosystem, which is described as dominated by a publish or perish culture leading to time as well as workload pressures (Sørensen et al., 2026), that the lab climate can only alleviate to a limited extent. Furthermore, the relationship between a good lab climate and research integrity is complexified and certain dilemmas are being introduced between good social relationships within the lab group and considerations of research integrity, as collegiality and friendships can conflict with rules and norms of proper conduct, like authorship allocation.

Ultimately the researchers' association of 'good research' with the ethicality of their conduct in interactions with colleagues and participants presents 'caring for good science' as the intricate and at times ambivalent process of 'working towards a socio-material collaboration' (Brenninkmeijer & Pols, 2026, p. 69).

Chapter 7: Preregistration in the lab

In this chapter I outline the researchers' discourse about how they enact preregistration as a novel reporting standard as a combination of credibility and hardship. I first provide some background about the emergence of and motivation behind preregistration in the biomedical, cognitive, and social sciences, followed by some critical reflections on this practice. Afterwards I review some of the existing research into the practice of preregistration. To address research sub-question two concerning the researchers' motivations behind practices that have been linked to research integrity, in the empirical analysis, I will first go through the motivations behind practicing preregistration in terms of what it does for them and their research (*utility* see Gläser, 2024). Specifically, I will show how the researchers present this procedure as aiding them in the conduct and communication of a well-thought-out study, reining in researcher degrees of freedom as well as their exploitation, and clearly distinguishing between the confirmatory and exploratory components of their research. I argue that all of these can be subsumed under the overarching aim of using preregistration as a reporting tool to manage credibility within the research community one is dependent on as a scholar (Aradau & Huysmans, 2019; Daston, 1995; Daston & Galison, 1992; Latour & Woolgar, 1979; Meskus et al., 2018; Shapin, 1984, 1995a; Shapin & Schaffer, 1985). Subsequently, I will present the practicalities and hardship associated with the practice of preregistration as discussed by the researchers. I will go through their stories of uncertainty surrounding preregistration in the context of piloting, the required investment of time and workload to use preregistration as well as how such necessary investment can be alleviated, how preregistration might be experienced like a prison or even as a punishment, and lastly how preregistration might overburden and exploit students. To conclude, I will briefly summarise the experiences and perceptions of the researchers and then explore what they might mean in the context of existing literature and research into preregistration as well as classical science studies literature about how researchers acquire credibility in their specific scientific community and the larger transepistemic arena to position them in the current research integrity discourse (Daston, 1995; Daston & Galison, 1992; Knorr-Cetina, 1982; Latour & Woolgar, 1979; Meskus et al., 2018; Shapin, 1984, 1995a; Shapin & Schaffer, 1985).

Introduction

Following the aforementioned claims of a credibility crisis, reformers, and especially those originating from psychology (social psychology) and biomedical fields, proposed and started to implement interventions to improve science. In the social, cognitive, and medical sciences, preregistration has recently been one of the most prominent proposed procedural interventions for improving research integrity (Lakens, 2019; Lakens et al., 2024; Nosek et al., 2018). Sometimes it is also called a pre-analysis plan (PAP, see e.g., Ofosu & Posner, 2023), henceforth, for simplicity, I will call it preregistration. This proposed intervention quickly started to spread across disciplines, especially in the social sciences. Similar to Vazire (2018), who talked about a 'credibility revolution' (p. 411) ushered in by the Open Science movement, with preregistration and practices like direct replication as crucial components, Nosek and coauthors (2018), focusing solely on the supposed impact and virtue of preregistration, proclaimed a 'preregistration revolution' (p. 2600). The idea of preregistering research follows the longer history of registration processes and repositories for clinical trials in medicine. There members of the International Committee of Medical Journal Editors (ICMJE) came together to make registration of study protocols mandatory in 2005 as a precondition for researchers to publish in their journals (De Angelis et al., 2004). In simple terms, preregistration is the procedure of specifying the hypotheses, as well as the research and analysis plans prior to the actual study and then depositing said plan in a preferably independent repository for others to access. Preregistration can be done with different levels of detail, which in turn requires varying amounts of work for preparation and execution or appraisal. Platforms (repositories) where researchers can preregister their research are, for instance, AsPredicted.org and the Open Science Framework (OSF). On the OSF, for example, there are currently 11 different preregistration templates, such as replication recipe (Brandt et al., 2014), preregistration for qualitative research (Haven & van Grootel, 2019), and systematic reviews (COS, 2025). Some scholars restrict the applicability of preregistration to certain types of research, arguing that it provides the transparency needed for the broader research community to assess the severity of statistical tests — how well a test would have detected an error if the claim were false — within a falsificationist framework of knowledge production (Lakens, 2019; Lakens et al., 2024).

More broadly, preregistration has been linked to various epistemic functions, typically framed as responses to the diagnoses underlying the crisis proclamations aiming to improve science (Nelson et al., 2021; Nosek et al., 2018; Romero, 2019). A central aim is to identify, control, and reduce questionable research practices (QRPs), particularly noted in quantitative

psychology and biomedicine, such as HARKing (hypothesising after the results are known), p-hacking, and selective reporting (Bergkvist, 2020; Kerr, 1998; Simmons et al., 2011). By attempting to restrict 'analytical flexibility' and 'researcher degrees of freedom' — seen as key enablers of QRPs — preregistration is hoped to reduce opportunities for exploitation and fishing expeditions for statistical significance, thereby lowering the probability of false positives (Simmons et al., 2011; Wagenmakers et al., 2012). Related to the goal of identifying and eliminating QRPs is also one of the broader purposes of Open Science reform more generally, namely, to identify, control or even prevent bias in research, such as confirmation bias and hindsight bias, in order to safeguard the impartiality and objectivity of the researcher and thereby the purity and integrity of the research processes (Daston, 1995; Derksen & Field, 2021; Flis, 2019; Hardwicke & Wagenmakers, 2023; Morawski. 2020; Penders, 2022). Furthermore, preregistration is intended to provide a transparent and time-stamped indicator that clearly separates exploratory research (postdiction) from confirmatory research (prediction) by establishing the temporal order of analysis, and separating hypothesis formation from testing (Nosek et al., 2018). Some appear to prefer confirmatory inquiry as more valuable or scientific than exploratory research, as articulated in the 'Agenda for Purely Confirmatory Research' (Wagenmakers et al., 2012, p. 632). Although, some also argue for the preregistration of exploratory research (Dirnagl, 2020). Importantly, because of its link to the replication crisis diagnosis, the presented ultimate aim of preregistration is to increase the transparency, replicability, and credibility of (psychological) science (Nosek & Lindsay, 2018). Moreover, preregistration is part of a broader cultural shift towards open and reproducible research promoted by the Open Science reform and by stakeholders such as the Center for Open Science (Nosek, 2019; Vincent & Hooper, 2025). To that end incentives, such as direct payment in the form of preregistration prices from the Center for Open Science, or requirements at journals (e.g., level three in the TOP guidelines), have been introduced to steer researchers towards preregistration (Nosek et al.,2015; Grant et al., 2025). Additionally, some journals employ conspicuous badges that advertise the value of preregistration and signal to the community that the reported research was preregistered, indicating that the preregistration can be consulted, presumably in the hope that such markers will convey an appearance of greater credibility (Klonsky, 2025).

There is also an extension of preregistration called registered reports (Chambers, 2013). Registered reports comprise of two stages. The first stage is like a preregistration, one writes down the study design and analysis plan, but unlike a preregistration this is then submitted to a journal for peer review. In case the submission is in-principle accepted, meaning irrespective of the

results the study will be published, if the submitted protocol is followed, afterwards, researchers collect the data, do the analysis and have to report the deviations from the protocol as well as potentially justify them. Lastly, in stage two researchers submit the full manuscript with the results (as well as discussion and conclusion). Now reviewers assess the manuscript and if the preregistration was followed and deviations are deemed justifiable the manuscript will be published (see e.g., Lakens et al., 2024). With registered reports the peer review process is moved prior to the production of the results, therefore, publication bias based on a preference for certain types of outcomes, like statistical significance, can in principle be reduced (Lakens et al., 2024; Rosenthal, 1979). Moreover, moving the review process prior to the actual conduct, unlike usual journal peer review, registered reports do have the potential to not just affect the reporting of research, but can actually influence the knowledge production processes of studies.

Critical perspectives on the practice and valuation of preregistration

Preregistration, both as a practice and as a regulatory instrument within the reward structure of science, is hardly uncontroversial. To the contrary, it has garnered considerable criticism. Preregistration is seen as promoted as a practice for 'better' or 'good science', framed within a narrative of a broken science in crisis (Bartscherer & Reinhart, 2025; Peterson & Panofsky, 2023), supported by what some critics regard as misguided and simplistic overly generalising heuristics, for instance, that QRPs are highly prevalent and problematic, that exploratory research tends to be inherently more speculative than confirmatory research, and that there is a general lack of replicability or reproducibility across different disciplines (see e.g., Rubin, 2023). This is seen as problematic because replicability or reproducibility are assumed to be near-universal epistemic practices and values for proper conduct across the sciences, despite contrary arguments from science studies scholars and researchers from within various disciplines (see e.g., Haig, 2022; Patton, 2024; Penders et al., 2019, 2020; Rubin, 2023; Ulpts et al., 2025). While proponents within the Open Science movement tend to conceptualise preregistration as part of a 'method assemblage' as Law (2004) calls it that protects and improves research integrity (Derksen & Morawski, 2022, p. 1497), more critical voices describe it as a 'superficial veneer of rigor' (Devezer et al., 2021, p. 16). In line with this critique, Devezer and Penders (2023) warn against treating preregistration as a marker of research quality, credibility or evidential value. Their concern is illustrated by Folk and Dunn (2023), who excluded non-preregistered studies from a systematic review on grounds that such work had

low evidential value. According to Devezer and Penders (2023) this kind of mechanistic use of preregistration where publications are filtered out solely based on preregistration status, implies that actual methods and metrics of excluded publications are not assessed. They caution that such practices could obfuscate the literature by obliviating certain publications, especially given the crucial role of literature reviews in shaping citation patterns as the collective memory of certain epistemic communities (Devezer & Penders, 2023).

One critique is that the sharp distinction between exploratory and confirmatory research, which underlies the motivation for preregistration, is problematic or unrealistic in light of statistical theory, philosophy of science, and the actual nature of scientific practice, particularly in the social and behavioural sciences (see e.g., Devezer et al., 2021; Rubin, 2022; Rubin & Donkin, 2024; Szollosi & Donkin, 2021). Along similar lines the apparent prioritisation of confirmatory research as epistemically superior to exploratory research has been linked to and criticised as a form of predictivism (Choi, 2024). Another critique is that preregistration, as a procedural and bureaucratic intervention, reflects the strict methodological focus of the current Open Science reform movement – methodologism – that distracts from other potentially crucial issues like the role of theory and theoretical deficits in some social and behavioural sciences (Devezer et al., 2021; Oberauer & Lewandowsky, 2019; Rubin & Donkin, 2024; van Rooij, 2019; Szollosi et al., 2020). Especially since some argue that preregistration is redundant, because if one has a well specified and comprehensive theory justifiable ways to conduct a study would be natural derivations from those (Flis, 2022; Szollosi et al., 2020).

Moreover, due to the aforementioned variation in detail and the complexity of research designs and data structures, preregistration can become a burden, creating additional workload for researchers, editors, and reviewers (Hostler, 2023; Ofosu & Posner, 2023). This problem is exacerbated by the need for some monitoring body to ensure that preregistrations are done properly and that the analyses are conducted in accordance with the preregistered details (see e.g., McDermot, 2022). Just having preregistrations lying on some repository without someone looking at or checking them, does not add anything to the control or self-correction mechanism of science. Additionally, some have identified a questionable research practice related to preregistration known as PARKing (preregistering after the results are known), in reference to the QRP of HARKing (Yamada, 2018). An example of PARKing would be when researchers conduct several (slightly) varied studies in advance to determine the likely outcome and then preregister and publish the version that produces the desired results. (Klonsky, 2025; McDermot, 2022; Pham & Oh, 2021; Yamada, 2018). This can, for instance, be done through pilot studies, prior to the official study that is then preregistered based on the knowledge

gained. Critics argue that this practice risks perpetuating or even worsening inequalities in academia by privileging large, better-funded labs that can afford extensive piloting to ensure their preregistered studies work as intended and capitalise on new standards and criteria that revolve around preregistration and their adherence (McDermot, 2022).

Klonsky (2025) explains inappropriate uses of preregistration in the context of a science system with Campbell's law, as preregistration is incentivised or even mandated and thus perceived as an indicator for 'good science'. Klonsky (2025) warns that psychological science already has a problematic history of elevating specific methods, such as statistical significance testing, as markers for proper science, a tendency Danziger termed the 'methodological imperative' (Danziger, 1985, p. 1). From this perspective the incentivisation and mandating of preregistration may simply be a case of history repeating itself. As Campbell (1979, p. 85) observed: 'The more any quantitative social indicator is used for social decision-making, the more subject it will be to corruption pressures and the more apt it will be to distort and corrupt the social processes it is intended to monitor.' Since, in some communities, preregistration is now widely used as an indicator for 'good science' (see e.g., Cadwallader et al., 2023; Devezer & Penders, 2023), the practice itself risks distortion, and its intended purpose may be undermined in the process.

Researching preregistration

As the practice and appraisal of preregistration is a somewhat recent development (besides clinical trial registration), research on it remains scarce, whether as a practice or an epistemic norm for credible research. The existing literature indicates that preregistration, if done in proper detail, can in principle facilitate the assessment of research (see e.g., Lakens et al., 2024). However, when we look at the research conducted on the actual practice of preregistration as currently implemented, we find some nuance and complications. Sarafoglou and colleagues (2022) surveyed researchers' perceptions of how preregistration affects their research workflows. A caveat, considering the representativeness of the survey, is that the respondents were mainly from the psychological sciences. The results from the survey suggest that researchers experience or expect preregistering a study to have a positive influence on the design of the study, the hypotheses, and on the analysis plan (also see Toth et al., 2021). However, the responses to the survey also indicate that for the researchers, preregistration has or is expected to have adverse effects on their work-related stress experience (also see Toth et al., 2021). This is in accordance with Ofosu and Posner's (2023) research indicating that writing a preregistration can be associated with a significant workload (see Hostler, 2023).

Moreover, van den Akker and co-authors (2024b) compared a corpus of preregistered articles with a corpus of non-preregistered studies selected from Web of Science and matched by topic. They set out to test five hypotheses commonly represented in Open Science reformers', and research integrity scholars' discourse about what preregistration can do or does for the research (van den Akker et al., 2024b): H1) the preregistered studies report a lower proportion of positive results than non-preregistered (which could work against skewed literature and publication bias), H2) the preregistered studies report lower effect sizes than non-preregistered studies, H3) the preregistered reports contain less statistical inconsistencies than non-preregistered studies, H4) more preregistered studies report power analysis than non-preregistered studies, and H5) the sample size of preregistered studies is larger than in non-preregistered studies. Van den Akker and colleagues (2024b) were unable to find evidence supporting hypotheses one to three and only found evidence for hypotheses four and five. Furthermore, van den Akker and colleagues (2023) also compared preregistrations with the actual corresponding published papers in psychology and concluded, in line with worries about lack of checking, lax application, and distorted preregistration practices (see e.g., Klonsky, 2025), that hypotheses reporting in the actual publication often deviates from the preregistration (also see Claesen et al., 2021).

To expand on current research into preregistration practices and their effect on the scientific literature, van den Akker and others (2024a) introduced the concepts of 'producibility' of preregistration, meaning whether the information provided in a preregistration is detailed and comprehensive enough to allow for the implementation in a study (also see Ofosu & Posner, 2023), and the 'consistency' of preregistration, which is the degree to which the actual publication corresponds to the details in the preregistration (van den Akker et al., 2024a). According to van den Akker and co-authors (2024a), both producibility and consistency must be ensured for preregistration to work against QRPs. Therefore, they also introduce the notion of preregistration 'effectiveness', which is a combination of the degree of producibility and consistency of preregistration. Research indicates that the preregistration templates on OSF work better in regard to producibility and effectiveness than AsPredicted.org (van den Akker et al., 2024a). This should not be too surprising considering that OSF allows for more extensive preregistrations than AsPredicted.org. Nevertheless, their research also indicates, that the performance of preregistration does not improve within their investigated timeframe (van den Akker et al., 2024a). However, one caveat considering the representativeness of van den Akker and colleagues (2023, 2024a, 2024b) studies, is that the investigated papers were predominantly from the psychological sciences. Furthermore, these findings of selective reporting despite preregistration should not

have been too surprising considering older research showing outcome reporting bias in registered clinical trials (Chan et al., 2004). Neither should it be surprising that the amount of detail contained in a preregistration plays a role. Brodeur and colleagues (2022, 2024) were also unable to find an effect of just preregistration on p-hacking in RCTs published in economics. Although, when controlling for whether the preregistration included a detailed pre-analysis plan, they found evidence that such detailed pre-analysis plans do reduce p-hacking. Moreover, neither of van den Akker and co-authors (2023, 2024a, 2024b) work does discuss or even refer to the relevant work of Brodeur and co-authors (2022) or Chan and others (2004).

Looking into the practices of editors and reviewers at PLOS Journals, Syed's (2025) findings 'indicate an extremely low rate of editors and reviewers engaging with preregistration plans during the review process' (p. 7). Which is problematic according to Syed (2025), because the intended function of preregistration to identify and prevent QRPs, resulting from high degrees of analytical flexibility, seems only achievable if there is actually a process of checking or monitoring implemented as well. There is not much research available on registered reports. Soderberg et al. (2021) let researchers do toy peer reviews of psychology and neuroscience articles based on 19 quality criteria, half of which were registered reports and the others traditional publications. The registered reports publications largely outperformed the traditional publication in the reviewer quality criteria ratings (Soderberg et al., 2021). Although the 19 criteria have such a large conceptual overlap, that it is doubtful whether they really measure researchers' perceptions of credibility or quality of distinct characteristics of the research. For example, two of the criteria are the rigor of the methodology and the quality of the methodology. Begging the question whether such conflated constructs lead to an artificial impression of increased quality or credibility ratings.

Field and co-authors (2020) attempted to investigate the effect of preregistration on researchers' trust in findings. However, while the authors remain convinced of the virtues of preregistration and registered reports when it comes to the quality and the trustworthiness of the research, their results were ambiguous and the study was unable to find evidence for or against trust inducing effects of preregistration. Moreover, Field and colleagues (2020) praise the role of registered reports in uncovering and presenting the at times messy nature of practicing research. Conry-Murray and co-authors (2022) found slight evidence for a positive effect of preregistration on researchers' judgement of abstract's trustworthiness. Hence the actual practice of preregistration still remains under-researched, and it remains an open question why researchers engage with the practice of preregistration or not, and how it influences their workflows, especially outside of the (strictly) psychological

domain. Therefore, in this chapter I will examine how the researchers in the labs I observed talked about preregistration.

Preregistration talk in the lab

During the field observations, I noticed that the topic of preregistration kept coming up in conversations between researchers, especially in supervision meetings where supervisors often asked whether specific studies were preregistered. An example from such inquiries by the supervisors is this excerpt from my fieldnotes during one of the supervision meetings in the cognitive science lab in the UK: The PI asks whether the postdoc has preregistered to which the postdoc answers no, it is a pilot. Therefore, in the interviews, I inquired into the motivations and potential utility of preregistration for the researchers (Gläser, 2024; Latour & Woolgar, 1979).

Thinking through your study with preregistration

Similar to the results from Sarafoglou and colleagues' (2022) survey, many researchers talk about how preregistration can help with planning and thinking through one's research to ensure a more thought-out study design (also see Spitzer & Mueller, 2023). In that sense it is described as a tool that somehow enables approaching the research more carefully from the beginning by thinking through important decisions like what are the hypotheses, methods, measurements, and analyses that are consequential throughout a study. Likewise, preregistration is also described as a procedure that makes things more specific and targeted by helping in ordering or structuring one's thought as well as the research.

> It is a good idea to think carefully beforehand about which analysis you want to do, which methods do you want to use, […] how you measure, how many participants you need, how many trials per condition, how many blocks, where you have more sessions[41] (Postdoc Neuroimaging Lab, Germany).

> It forces you to think ahead and also to really structure your research. […] It helps you even in those early stages to make sure that what you're researching is what you want to research and that you will be able to test the things you're interested in (Postdoc Neuroimaging Lab 1, The Netherlands).

[41] Original in German: ‚Grundsätzlich finde Ich das schon eine gute Idee also, dass man sich wirklich vorher genau überlegt welche Analysen willst du machen, welche Methoden willst du benutzen. Davon hängt natürlich auch ab wie du misst, wie viele Probanden du brauchst, wie viele Trials pro Bedingung, wie viele Blöcke, ob man überhaupt mehrere Sessions hat oder irgendwie sowas. Solche Fragen die kannst du eigentlich nur richtig beantworten eigentlich wenn du weißt was du nachher mit den Daten machen willst. Und das sich vorher genau zu überlegen ist schon grundsätzlich eine schöne Idee.'

One scholar even goes further into detail and directly compares the advantage or function of preregistration as a thinking-aid in planning towards a well-thought-out study as well as explicating one's hypothesis to other aims that are often presented in the literature as improvements of science provided by preregistration. Among these other claimed utilities of preregistration are the identification or prevention of QRPs like HARKing and p-hacking (Lakens et al., 2024). Interestingly, the researcher judges the potential function of preregistration as a thinking and planning aid through explication as more valuable or relevant than its often-claimed utility as a check for QRPs (Bergkvist, 2020). It seems to be less about prevention of potentially problematic behaviour and controlling oneself or others, but in a rather constructive sense, it is more about using it as a procedure to structure one's thought and properly plan the research. It is here seen then as a practice that can thereby help to make a study more powerful in the process of knowledge generation in a falsificationist framework, reminiscent of Lakens and colleagues' (2024) restriction of preregistration's applicability to check or improve the severity of hypothesis tests in a falsificationist knowledge production framework.

> It typically improves the research, because it forces the researchers to be more explicit about the hypothesis that they have and how their experiments will actually contribute to knowledge by sort of falsifying or not falsifying the hypothesis. I find that a much more important aspect of preregistration, than this trying to not hypothesise after results are known (HARKing) or to make sure that people are not continuously testing whether there's something in their data and then stopping when they have any significant effect (PI Neuroimaging Lab 1, The Netherlands).

Preregistration to constrain researcher degrees of freedom

An issue that comes up in the literature as one of the most popular narratives in the justification of the need for preregistration as a procedure to check or control researchers, in the context of problems in science that have been identified in association with claims about irreplicability and a lack of credibility of research findings (especially in the social sciences), are so-called 'researcher degrees of freedom' (Simmons et al., 2011, p. 1359) or 'flexibility' (Ioannidis, 2005, p. 698) during various stages of a study, where researchers have to navigate through, what Gelman and Loken (2014, p. 460) called, the 'garden of forking paths' (also see e.g., Nosek et al., 2018). This roughly means the number of decisions during various stages of the research that scientists can make, justifiably or not (see also van Drimmelen et al., 2024). Such decisions are seen as highly important because they can heavily influence the results. To be more precise, the assumption is that it is the combination of degrees of

freedom, the sensitivity of the statistical results to said degrees of freedom and the biased human nature, that will lead researchers to also recognise statistical flukes as significant positive results, which has a detrimental effect on the scientific literature (Flis, 2022). Therefore, this freedom is also seen a potentially problematic, because it can be exploited on the hunt for statistical significance, especially in a context where that is rewarded (see e.g., Simmons et al., 2011). In this context interlocutors describe preregistration, for instance, as a tool to communicate to the research community that these degrees of freedom have not been exploited through 'some kind of weird arrangements' (PhD Neuroimaging Lab 2, the Netherlands) and it thereby seems to be further used as a bureaucratic or procedural indicator for research quality, credibility or integrity, by making the research processes more visible and potentially auditable (Penders, 2022; Shapin, 1984). Hence, the practice of preregistration seems to be used in a process of indicating proper statistical conduct to a research community that is suspicious of certain practices like claiming statistical significance based on p-value smaller than 0.05.

> That you are able to tell other researchers, look here I have registered exactly that analysis already beforehand. […] then you might somehow have more trust, or the people have more trust in your results, as when they have no idea what you can do. With fMRI and things like that the degrees of freedom are extremely high […] and then you can, of course, quickly go on a fishing trip and somehow try 30 things and one of them is coming out significant and that is the one you report. […] And you don't know that when there was no preregistration beforehand[42] (Postdoc Neuroimaging Lab, Germany).

[42] Original in German: ‚dass du anderen Researchern sagen kannst kuck mal ich habe genau diese Analyse auch vorher schon angemeldet sozusagen, dass Ich das so machen werde und nicht anders und das hat natürlich dann hast du vielleicht irgendwie mehr Vertrauen oder die Leute haben Vertrauen in deine Ergebnisse als wenn Sie keine Ahnung haben was du machen kannst. Bei fMRT und so sind natürlich die Freiheitsgrade extrem hoch. […] und dann kann man natürlich schnell auf einen Fishingtrip gehen und irgendwie 30 Sachen ausprobieren und eins davon wird signifikant und das reportest du dann. […] Und das weiß man natürlich nicht wenn es keine Preregistration vorher gab. Dann weißt du nicht was die Leute alles ausprobiert haben. Und in der Weise find Ich das schon ganz schön, das zu machen wenn es geht. […] Also es sind immer irgendwelche arbiträren Entscheidungen. Aber das ist natürlich auch wieder cool, wenn man das Preregistriert und vorher schon sagt bevor ich meine Daten angekuckt habe. […] dann kann jemand das nachschauen und sagen okay das hat der auch vorher gesagt dass er das so machen wird. Und dann kann man zumindest diesen degrees of freedom nicht mehr durch tausend, theoretisch durch tausend, p-values austesten könntest oder so. Unsere Ergebnisse hängen superstark davon ab wie du sie analysierst. Also dieselben Daten du kannst ein völlig anderes Ergebnis rauskriegen […] Und deshalb sind solche Entscheidungen extrem wichtig, vor allem statisches veranschaulichen nachher.'

> It makes me feel better, because I know I'm holding myself onto some pre-made plan. So, you know, people are very biased [...] And if you actually write these things down, put it somewhere on the internet, then you can trust yourself more. [...] That study, which I first did as a registered report. [...] We did the analysis, and the main effects, P value, [...], was 0.04, so one point below the significance level, which looks like P hacking. [...]. When I saw it, I was like, that's bad. But then I realized that I did what I set out to do. [...]. And since that day I sleep better because I did not do the p-hacking, I had no chance of doing it (Postdoc Neuroimaging Lab 1, The Netherlands).

This researcher describes how preregistration or registered reports make one even trust oneself more, 'feel better' and 'sleep better'. This response also makes again clear that there is a perception that the community is showing something like a culture of suspicion towards a specific type of results, because they are seen as indicators for potential misbehaviour or exploitation of the before mentioned flexibility and degrees of freedom. The assumption seems to be that researchers cannot be trusted (Flis, 2019; Nuzzo, 2015). This is linked to the QRP of p-hacking with p-values just below 0.05 being seen a suspicious, because they could indicate that someone stopped testing once that value was below the standard significance level of 0.05. The researcher, on the one hand, talks about preregistration holding oneself accountable by keeping the research to a predefined plan, but on the other hand, also mentions the role preregistration plays during the social process of communicating one's results to the research community (Shapin, 1995a). Therefore, preregistration, and registered reports even more so, seem to be seen as mediums that communicate that no such misbehaviour has taken place. In other words, it is to a certain degree also seen as a safety net or defence against suspicion from the research community as well as oneself. Hence, one could argue that it is a credibility management strategy within a culture of suspicion, at least surrounding certain statistical practices, that could aid researchers in the social process of persuading the community of one's and one's research credibility by making them 'virtual witnesses' (see e.g., Meskus et al., 2018; Shapin, 1984, 1995a; Shapin & Schaffer, 1985).

Practicing preregistration seems to not just be motivated by researchers expecting some type of suspicion from their epistemic community, but also by them being suspicious of themselves when it comes to the conduct of their research. This is somewhat reminiscent of Penders (2022) use of 'Fremdzwang' (pressure from others) and 'Selbstzwang' (pressure from oneself) in the context of conceptualising the bureaucratic and procedural interventions of preregistration and registered reports as part of a civilising process, which is the current OS reform movement. This perception of suspicion that is directed inwards and outwards, seems to come from a conviction that humans

are inherently biased (Flis, 2019; Morawski, 2020), as the scholar put it: 'people are very biased' (Postdoc Neuroimaging Lab, the Netherlands). This suggests that, in line with Daston and Galison's (1992) notion of nonintervention-ist objectivity, there is a conviction that due to the combination of degrees of freedom and researchers' bias the research processes need to be protected from the scholars' intervention (Field & Derksen, 2021; Penders, 2022). Here such protection is apparently provided by the procedures of preregistration and especially registered reports, through the literal explication and sharing of research processes which is assumed to provide the possibility for control and correction (Flis, 2022).

In the literature preregistration is proposed as one of the interventions to address or even solve the replication crisis that is claimed to afflict some scientific disciplines (see e.g., Resnik & Shamoo, 2017; Peterson & Panofksy, 2023; Romero, 2019). Interestingly, this next researcher explicitly links the procedure of preregistration to an apparent irreplicability issue in psychology, when asked for the reason to apply the practice. This not just indicates that the perception or narrative of a crisis state concerning replicability or credibility has reached or is shared by some of the researchers in the psychology community, but also that such a perception of a crisis seems to play a role in why researchers use preregistration. 'Because there was historically a lot of problems with having effects in psychology that simply didn't replicate' (Postdoc neuroimaging Lab 1, The Netherlands).

Thus, while such crisis perception or narrative seems to fuel a culture of suspicion, in the context of these degrees of freedom or 'wealth of possibilities', preregistration is seen as a form of scientific reporting that keeps you as a researcher honest or enables you to communicate honestly to the research community in the face of the temptation presented by bias combined with degrees of freedom. Here preregistration is described as an intervention that should persuade the community that the reported finding is not the consequence of the analytical flexibility inherent to the research, but a consequence of an a priori set path. Hence, here again we see the aspect of communicating credibility (Shapin, 1995a). Moreover, it is further described as a way of putting the value of honesty, which can be found in most conduct guidelines (Desmond & Dierickx, 2021), as a form of transparent reporting into practice: 'Given the wealth of possibilities of what to do with the data also to be able to be honest and say like, this is what I'm interested in, this is the critical comparison' (Postdoc neuroimaging Lab 1, The Netherlands).

A timestamped firewall between confirmation and exploration

As the preceding quotes and wordings such as 'previously' and 'beforehand' suggest, according to the scholars an important circumstance in regard to the

credibility of the reported research seems to be time, or the sequence of events leading up to the reported findings and especially 'when' certain decisions pertaining to the research, analyses, and hypotheses were made. As the following two quotes further show, we again see a combination between inwards and outwards looking motivations for preregistration (Penders, 2022). As such preregistration is supposed to provide a paper trail enabling scholars to communicate the order of events in the research.

> For other people to show that your hypothesis is the one you said before you collected data. I think it's also for yourself. You really settle on a hypothesis and not later on, you feel like, yeah, actually maybe it's different. I mean, of course that can happen, but I think it's important that you first put it there. That was your initial idea and then later it wasn't right but then feel free also to share it with the community, I did this, but we found this. And that's why my idea developed like this (PhD Neuroimaging Lab 2, The Netherlands).

> Because for myself also, it's very easy to forget when you've decided, oh, it didn't work. No, of course it didn't work. Now let's have a look at the data and you go like, oh, but this makes much more sense, [...] And it's very easy to then actually rewrite the whole thing. And say, this is what the prior literature said. That's what the hypothesis is. And so now you've, essentially been overly flexible. [...] You have to tell the audience, your community, why you set up the experiment, what you predicted and then that you didn't find it. But there's something else. And then you name and still write that in a way that you share the message properly. That your, communication is effective (PI Neuroimaging Lab 2, The Netherlands).

The last part of this previous quote also suggests a fear that the wealth of information shared to ensure a clear distinction when decisions during the research were made in the name of communicating credibility, can potentially hide the actual message that a research paper is supposed to convey to the epistemic community. One could ask whether transparency procedures of scientific reporting, like preregistration, could actually lead to a situation where important information or messages are hidden in an abundance of information in the name of Open Science (Elliot, 2022)? In a more pessimistic view, akin to document dumps in legal contexts, one might be tempted to ask whether transparency can also be strategically used for the purpose of obfuscation instead of clarity.

In the context of scientific research as a sequence of practices in a specific order, some researchers also explicitly touch on the textbook distinction between confirmatory and exploratory research and what role preregistration plays in enabling oneself to uphold and transparently communicate this distinction to the wider community, as a sign of credibility (see e.g., Nosek et al., 2018). Therefore, an important function of preregistration is to demonstrate

that the hypothesis was specified prior to data collection, effectively serving as a timestamped firewall between exploration and confirmation, signalling that the hypothesis was not shaped by the data it was tested against (Nosek et al., 2018).

> To communicate to the outside community, what your hypothesis beforehand were and what exploratory analyses were that came up during it. So, in this way it is both for ourselves it is important, but also it is just a transparent research process to the outside and many journals even require it. It is not just important but also necessary (PhD Neuroimaging Lab 2, The Netherlands).

However, in the end the scholar also mentions that using preregistration is not just intrinsically motivated by the wish to improve the credibility of one's research in the context of the norms and values of the research community, but also extrinsically through force by journal publication requirements (Grant et al., 2025). This for instance is the realisation of the last stage of Nosek's (2019) culture change pyramid going from making open science practices possible and easy over making them normative and rewarding to finally making them required (also see Nosek et al., 2015; Ulpts et al., 2025; Vincent & Hooper, 2025).

During the interview a PhD supervisor and postdoc in a German psychology lab, who are also themselves involved in metascience and the larger Open Science reform movement, talks about a discussion with a PhD student who wanted to preregister their work. However, in this instance the supervisor voiced arguments against this procedure which the supervisor usually supports. The postdoc judged preregistration as unfitting because, on the one hand, the study in question was exploratory and, on the other, one apparently has to be strategic about where to invest the effort to do scientifically 'good research', hinting towards what has been called perverse incentives (see e.g., Benedictus et al., 2016; Edwards & Roy, 2017; Robinson-Garcia et al., 2023; Vendé et al., 2025) as well as previous research on researchers' perception of research integrity that there can be a conflict between the researchers' own notions of 'good science' and the institutionalised criteria for career advancement (see e.g., Brenninkmeijer & Pols, 2026; Davies, 2019; Elizondo & Kaltenbrunner, 2024; Hangel & Schickore, 2017; Pols et al., 2024; Schickore & Hangel, 2019). Hence, at least in their attempt to justify or explain themselves to me, the scholars do show something like an investment strategy, concerning the expected payoff and the costs of practices that are associated to the credibility of research (Latour & Woolgar, 1979; Peterson & Panofsky, 2021; Shapin, 1995a):

> So, [PhD], for example, wanted to preregister their [...] study, and I said, well I am a fan of preregistration, but I know for myself that this project is a rather

> exploratory project and we wanted to report it like that. However, [PhD] still really wanted to and also wrote something down then. But that was something where I spoke against a practice that I also often advocate for, that happens too. Yes, but I think one also has to be careful, so you don't get so far. […] one does not stay in academia if you don't think strategically somehow. If you don't think at all at some point, this project is better for my career. Thus, if you completely ignore such things and is only guided by one's own interest and what is scientifically good, then I don't know how far one gets[43] (Postdoc Personality Psychology, Germany).

Another researcher shared the motivations behind using preregistration for their own research based on personal positive experience with preregistration and especially registered reports:

> My supervisor was very supportive, […] but also a bit dubious about whether all the time putting into it is going to be worth it. And by now I would say it actually is worth it because open science and people who do open science […] have some advantage in the job market (Postdoc Neuroimaging Lab 1, The Netherlands).

Here the Dutch scholar describes two systemic or external drivers of using preregistration. There is the supportive attitude of the supervisor, who while actually doubting whether the time and work invested in preregistration or registered report, are really worth the effort, was still facilitating practicing it. Furthermore, having a (potential) career advantage due to using Open Science practices like preregistration in the job market during research(er) evaluation is mentioned as making it worth the extra work, by now. Hence, the researcher seems to also indicate experiencing somewhat of a shift in the evaluation culture of academia more in alignment with Open Science ideals (see e.g., Gärtner et al., 2024; for evaluation regimes see Cramer & Reinhart, 2026; Robinson-Garcia et al., 2023). The here mentioned supportive attitude from this supervisor seems to be somewhat in contrast to the before mentioned more careful or pessimistic stance of the German supervisor who advised a PhD student to maybe not preregister exploratory research, because one has to be more strategic in the academic job market regarding where to invest the effort. Hence, there seem to be different perceptions, expectations and experiences

[43] Original in German: ‚So also [PhD] wollte zum Beispiel seine […]-Studie gerne präregistriert, und ich meinte, naja, ich bin Fan von Präregistrierungen, aber ich weiß jetzt selber, dass wir dieses Projekt einfach als, es ist eher ein exploratorisches Projekt und wir sollten das auch so berichten und irgendwie, aber [PhD] wollte es trotzdem sehr gerne machen und hat dann auch irgendwas geschrieben, so, aber es war halt, äh, das war für mich so ein bisschen sowas zum Beispiel, wo ich dann so mal gegen eine Praxis ausgesprochen habe, die ich auch schon oft selber befürwortet habe irgendwie, sowas kommt auch vor.'

coexisting about the research(er) evaluation system in Europe and how as well as to what degree Open Science and research integrity practices are rewarded.

> I try to preregister everything that I want to take to be what you might call a confirmatory analysis, right? Sometimes you're just running a pilot. I guess I do what I take to be the community standard, sometimes you're screwing around with, the data set just to understand what's going on, but you're not going to try to convince anyone of anything on the basis (Reader Cognitive Science, The UK).

Similar to the German postdoc, for this scholar preregistration is a procedure applicable to confirmatory research, which the researcher perceives as the community standard, something like normative pressure to preregister confirmatory research. In the end of that quote the interlocutor explicitly mentions the aspect of trying 'to convince' people, which again brings the social aspect of persuading other researchers or the wider community of their credibility (Gilbert, 1977; Shapin, 1995a), and preregistration being one reporting tool that can help in doing that for confirmatory research.

Another scholar also mentions the influence of the academic circles one is moving in when it comes to practicing preregistration, stating that 'It was the style of the time. Definitely simply because I was very active in the Open Science area' (Postdoc Personality Psychology, Germany). During the observations in a meeting researchers in a Dutch neuroimaging lab also talked about the role of local rules and norms in the lab, as the following excerpt from my fieldnotes shows:

> they are discussing how to address the issue of preregistration in the Lab handbook. Issues addressed are for instance, what is a preregistration, that preregistration is not mandatory yet, should preregistration be opt-out or opt-in and, how elaborate should a preregistration be.

While exploratory research is seen as somewhat ill-fitting for preregistration, because, for example, in the case of piloting sometimes researchers seem to just be 'screwing around' to acquire an understanding of how something actually works (Hangel & Schickore, 2017; Knorr, 1977; Knorr-Cetina, 1981; Schickore & Hangel, 2019). Similarly, another researcher lists some reasons for why not to preregister and when it makes sense to them. With the first reason being that the investment of putting in the work to preregister the study might not be worth it in cases where the researcher does not know whether the study is actually going to work. Hence, when researchers are just tinkering to get something to work (Knorr, 1977; Knorr-Cetina, 1981), preregistration is perceived as less valuable.

> The times when I don't is often when I'm not quite sure this study is going to work. I don't want to go through the full time of preregistering this study. Very

> often also it's because I don't really know what is the right statistical analysis for this design [...] I tend to prioritise preregistration when there's a key methodological point that I think might be contentious in the data analysis, for example, or in the way we implement a particular model (Postdoc Cognitive Science, The UK).

Hence, since researchers identify the sequence of events concerning hypothesis formation, data collection, and data analysis as crucial for the practice of preregistration, the research for which preregistration seems to be seen as most suitable is confirmatory and where they already know beforehand what will happen and that something works as intended.

Preregistration as a credibility management strategy

Considering the insights from the social studies of science, describing science as a fundamentally social endeavour with socio-epistemic processes of knowledge production in specific communities (see e.g., Bourdieu, 1975; Collins, 1985; Knorr-Cetina, 1999; Latour & Woolgar, 1979; Shapin, 1995a; Whitley, 1983), preregistration seems to be described as a procedural intervention for the social process of persuading the epistemic community one is a part of that their research is of high quality or integrity and therefore credible. Two of the characteristics that are part of the credibility message believed to be contained in preregistration are the before mentioned declaration of which parts of the research are exploratory and which are confirmatory as well as demonstrating that the existing degrees of freedom were not exploited despite human temptation and bias. This demonstration of credibility through preregistration as scientific reporting that attempts to make the community 'virtual witnesses' seems to exist within a culture of suspicion (Meskus et al., 2018; Shapin, 1984; Shapin & Schaffer, 1985). When talking about in which cases not to go through the whole time to do a full preregistration, these last quotes also hint towards a consideration that Latour and Woolgar (1979) in their conceptualisation of the cycle of credit, refer to as a consideration of the return on investment concerning the capital researchers acquire as scientists (Peterson & Panofsky, 2021). Here the form of credit is credibility in relation to the research community (Shapin, 1995a). Researchers also hint towards this role of preregistration to convince the epistemic community that one is credible, by stating that preregistration is used when something might be 'contentious', or they need to 'convince' the community of something.

> And lately, which is really nice, I think it actually makes you as a researcher more credible as well (Postdoc Neuroimaging Lab 1, The Netherlands).

> When you're in a position where you want to convince people that you didn't just exploit degrees of freedom to cook up an answer that you thought was made for a sexy story, that's when you want to pre-register as much as you can (Reader Cognitive Science, The UK).

However, one researcher also talks about that having 'the label preregistration' on their research can already function as an indicator for research quality to the community so they have more trust in you and what you did.

> To use the label preregistration later on, as a quality indicator of course[44] (Postdoc Neuroimaging Lab, Germany).

This phrasing of 'label' of preregistration and the indicated lack of actual inspection of preregistrations and comparisons with the published study (Syed, 2025), casts doubt on whether the practice of preregistration actually makes research more credible or is it just part of a social persuasion process based on what Devezer and colleagues (2021) call a 'superficial veneer of rigor' (p. 16), with preregistered studies being perceived within the community as more credible without actual assessment of said credibility (Devezer & Penders, 2023). Such worries are further supported by research into preregistration suggesting low rates of checking and high rates of suboptimal practices related to preregistration as well as lack of evidence for the claimed benefits of preregistration, like preventing QRPs such as HARKing (see e.g., van den Akker et al., 2024b; Syed, 2025). To paraphrase Devezer and colleagues (2021) this could potentially create a situation where incredible science is hidden behind a superficial veneer of credibility. Recent events surrounding the retraction of a paper by leading Open Science reformers proclaiming that rigour and research integrity enhancing Open Science practices like preregistration are working, also suggest a somewhat superficial use of preregistration without actual integrity improvements, because one of the reasons for the retraction were problematic and misleading preregistration practices (Bak-Coleman & Devezer, 2024; Protzko et al., 2024).

Exactly this frustration of preregistration being used as a superficial indicator for 'good research' and research integrity, actually came up during observations of a journal club meeting of the German personality psychology group members of a larger European project. Here is an excerpt from my fieldnotes of a conversation about a rejected manuscript that hints towards a form of publication bias based on whether a study is preregistered or not:

---

[44] Original in German: ‚das Label preregistration nachher zu nutzen, als natürlich Qualitätsausdruck, dass du anderen Researchern sagen kannst kuck mal ich habe genau diese Analyse auch vorher schon angemeldet sozusagen, dass Ich das so machen werde und nicht anders und das hat natürlich dann hast du vielleicht irgendwie mehr Vertrauen oder die Leute haben Vertrauen in deine Ergebnisse als wenn Sie keine Ahnung haben was du machen kannst.'

> They mentioned that it seemed as if it was rejected, because it was not preregistered and that they think that preregistration is sometimes used as a blind heuristic for 'good' or 'bad research' or in this case a accept or reject decision. It seemed, according to them, as if people were blindly in favour of preregistered studies without considering what preregistration was actually for.

I also talked with the German postdoc who mentioned using the label preregistration as a credibility indicator about whether the preregistrations are also checked during peer review to which the postdoc replied:

> I just believed that they actually did it that way. Although there was this paper recently where they said that even with preregistration, a huge number of people later deviate from what they were actually supposed to do… which isn't bad per se, but it's interesting that they don't stick to it[45] (Postdoc Neuroimaging Lab, Germany).

The first half of the quote indicates a somewhat superficial treatment and a lack of the assumed checking or control that is supposed to be provided by preregistration. This can be due to a lack of time in the context of the already stressful workload of most academics (Hostler, 2023). The second half of the last quote also indicates that the researcher actually shows an interest in research and publications concerning procedures like preregistration.[46] Since the UK lab also had job talks taking place while I was visiting them, the use of preregistration and Open Science badges for preregistration came up in that context. Here the issue of superficial use of preregistration is presented as a form of virtue signalling.

> I already see it in the job talks, people put the badges on every study, on the top of the slide. But nobody in the room is going to look at the preregistrations. And when I have looked at preregistrations as a reviewer, I think I've only done it once or twice. I've never seen them to be anything, but terrible. As in like, you say preregistered, I went over to the preregistration and it's unrecognisable to what you did. Or it's about the pilot study, but you made it sound like it was about all the studies. So, it's just, it feels like, oh, you just found a new thing to hide behind where it's like, trust me, bro, I've got a badge (PI Cognitive Science Lab, the UK).

This leads us to the next section about the researchers' experiences and perceptions with the practicalities and problems of practicing preregistration.

---

[45] Original in German: ‚Ich habe dann einfach geglaubt, dass die das tatsächlich so gemacht haben. Obwohl es gab, ja letztens ein Paper wo sie irgendwie gesagt haben wo sie gesagt haben, dass selbst bei preregistration extrem viele Leute nachher davon weg gehen von dem was sie eigentlich … was ja nicht per se schlimm ist, aber das ist natürlich interessant das man die (sich) dann doch nicht dran hält.'

[46] The paper that is referred to is probably from van den Akker and colleagues (2023).

Practicalities and problems with practicing preregistration

In the context of science reforms and regulatory attempts at securing and improving research integrity, preregistration can be treated as a new standard that is introduced into the practice of research. Therefore, the application of preregistration in research can be conceptualised as the enactment of a new standard in specific local sites of knowledge production (Millerand et al., 2013), with their unique and often idiosyncratic conditions. We do less see a resistance to the new standard of preregistration (Lampland & Star, 2009), but rather a story emerges between ongoing commitment to the (new) standard coupled with obstacles and difficulties with the local enactment of preregistration in practice (Millerand et al., 2013). Hence, when the researchers talk about their own experiences with preregistration certain considerations about practical and as problematic perceived aspects emerge.

The uncertainty of piloting

As previous quotes already indicate, the distinction between exploration and confirmation also seems to play a role in the context of piloting, because it reappears in discussions surrounding the question whether to preregister pilots and how to approach that. Pilots are described as a process of exploration towards figuring out what works for the research and what does not (Hangel & Schickore, 2017). Researchers seem to not just be unsure how to apply preregistration to pilots, but they also indicate a certain suspicion about what other researchers actually did and whether there was a pilot or not when someone indicates that their published research was preregistered. Therefore, this uncertainty associated with the practice of piloting in the context of preregistration could lead to a situation where preregistration is actually detrimental for transparency and credibility alike, instead of increasing it.

> I feel many preregistration things were done after a pilot study. So, you can see that from many papers they say they preregistered, and the results really replicate the preregistered hypothesis or everything. It's because they have done the pilot before. But they didn't say that in the paper. So, it feels very clean or everything but actually it's not. And they are a rich lab; they can do that. But they change the standard or the criteria of the whole area. And because for other labs, they don't have that much money to do that. So, their results, it's not as clear as those rich labs if they preregistered. They will always say that, ah, this is not consistent with our preregistration. I just have a 30 page of appendix to explain that. […] I guess if you have more money, you can run more experiments and you can really choose which to publish (Postdoc Cognitive Science, The UK).

This quote relates to the issue of PARKing in the context of pilot studies (McDermot, 2022; Yamada, 2018). In that context the scholar also hints towards a problem of inequality in the distribution of resources that are necessary to produce clean preregistrations and experiments in various amounts of numbers (Leonelli, 2022, 2023). The formulation of this researcher is further quite insightful, because it also relates to the quote from another researcher about preregistrations of direct replications being necessary, of high value, and unproblematic because everything is already known:

> It is completely unproblematic, absolutely necessary, but very easy is a preregistration, when it is about doing a replication. If I have a direct replication, yes logically, since I do it exactly like before. That's just simple and cheap. That is totally simple, there I can go into detail[47] (Professor Personality Psychology, Germany).

Hence, when preregistrations follow pilots, one could view the actual research that follows and is preregistered and ultimately published as a replication of the pilot study that worked. Hence, despite preregistration it can be seen as a type of invisible replication (Lubega et al., 2023; see Haig, 2022 for hidden forms of replication in science). Relevant is also the mentioning of rich labs being able to conduct multiple pilot studies and then preregister as well as publish the research that cleanly implements what the preregistration states (Hangel & Schickore, 2017). The researcher speculates that such behaviour in the context of incentivisation or demands for preregistration could lead to a system with unrealistic standards for preregistration, because the amount of funds and working power required for such extensive piloting is not equally distributed across labs or nations. Thus, such behaviour from well-funded labs could lead to a system wide expectation from, for instance, journals and reviewers that a good study is one where the reported research resembles the preregistered plan as closely as possible, without many or extensive deviations. This could also lead to a social pressure towards confirmatory research and against exploratory research, where 'clean' preregistration is part of an image of good scientific practice, required for appearing credible. In other words, this shift in standards related to preregistration could introduce its own kind of publication bias, where the criteria for publication become dependent on the practice of 'clean' preregistration. Hence, not just whether it is preregistered, but also to what degree the reported study follows the preregistered protocol can become a criterion for gatekeeping. Which is an even

[47] Original in German: ‚Es ist ganz unproblematisch, also absolut notwendig, aber ganz einfach ist eine Präregistrierung, wenn es darum geht, eine Replikation zu machen. Also wenn ich eine direkte Replikation habe, ja logisch, ich mache es nämlich genauso wie die vorher. Das ist ja einfach und billig. Das ist total einfach, das kann ich bis ins Detail gehen.'

stronger publication bias than the previously mentioned one where scholars in a journal clubs suspected that a manuscript was rejected just because it was not preregistered. Linked to this issue of how preregistration relates to pilot studies and how this could distort preregistration practices and expectations is a general vagueness or lack of guidance concerning preregistration in piloting.

> But I think that there's an awful lot of grey area there, because there's a sense in which almost everything starts out as a pilot. And it's very unclear whether one is supposed to preregister one's pilots when one has no idea what's going to happen yet (PI Cognitive Science, The UK).

Hence, this further indicates that while preregistration is supposed to, if not constrain, then at least make visible analytical flexibility and degrees of freedom in researchers decision making during various steps of a study by having to report the decisions they made during actual conduct of the study in comparison to a time-stamped plan of what they set out to do, it itself comes with certain degrees of freedom or, as the researcher called it, 'grey areas' that seem somewhat hidden (Klonsky, 2025; Yamada, 2018; also see van Drimmelen et al., 2024). Here the researchers describe the potential to exploit certain degrees of freedom when preregistering in the context of piloting. Apparently, because it is a grey area, researchers can preregister pilots or not and influence the process of preregistering the actual study that is eventually submitted to a journal and subsequently published. Therefore, some seem to experience a lack of guidance or control when it comes to preregistration in the context of piloting and how that relates to the nature of preregistration and the published manuscript. In other words, the researchers seem to feel that the practice of preregistration is undetermined, leading to flexibility and uncertainty in its application (van Drimmelen et al., 2024, 2025). These researchers describe it in a way suggesting, in line with McDermot (2022), that this leads to its own potentially questionable research practice called PARKing (also see Klonsky, 2025; Yamada, 2018).

Some of these quotes and the excerpt from my fieldnotes at the end of the last section align also with Klonsky's (2025) warning about how, in line with Campbell's law, preregistration becomes an indicator for 'good science' and then is corrupted over time in its application and reception. As the PI of the cognitive science lab in the UK, reinforcing the postdoc's comments, puts it when talking about the inherent gameability of Open Science interventions like preregistration:

> I guess my more cynical view is that all of these sort of new measures and prescriptions are incredibly gameable. Just as gameable, probably, as the statistical reporting of your way. […] But you can do all the same stuff in terms of

> generating multiple preregistrations or preregistering studies, but like not writing up all of the studies and things. So, many of the things feel a bit like goalposts moving if you think of scientists as bad actors (PI Cognitive Science, The UK).

Furthermore, the previously mentioned sentiment of not preregistering in cases when it is unclear to the researchers what works, comes up more often. This is again related to the textbook distinction between confirmatory and exploratory types of research that seems so fundamental to the idea(l) of preregistration (Nosek et al., 2018). This distinction becomes especially important or even somewhat questionable considering that some highlight the important or dominant role of exploratory components in their research. More specifically, according to some researchers there are inevitably unexpected aspects in the research that need to be dealt with and that research usually is a mixture of exploration and confirmation with some being more exploratory and others being more confirmatory, instead of a clear-cut binary distinction (Schickore & Hangel, 2019). Here, as a previous quote also already indicated, preregistration could then actually help in making the exploratory component of the day-to-day research more explicit to the outside.

> I remember someone doing a really horrendously long pharmacological imaging study, working on it for two and a half years. And then one day after de-blinding the drug codes they enter my office and say, it didn't work. […] Of course it doesn't work. It never works. The test is always different from what you predict. But that doesn't mean that you go like, okay, now it's done. Right? Now I'm done. It didn't work. I'm going to write it up. It didn't work. No, that's not the way our science works. It's very, very much exploratory. But we have to be much more explicit about that (PI Neuroimaging Lab 2, The Netherlands).

Hence, since their type of research is self-described as to some degree inherently exploratory, here preregistration is seen as a way to show that the research is way more exploratory than researchers often let the community know, instead of making their research more confirmatory. The reoccurring theme of having to know beforehand what will happen in the study when preregistering it, is also problematised as a factor that implies the danger of making the research that is produced under a norm of preregistration more generic. This could potentially work against an aspect of day-to-day research that scholars investigating good scientific practice, identified as a crucial component of 'good science', namely creativity and the drive to embark into the unknown territories of knowledge production (see e.g., Pols et al., 2024; Valkenburg et al., 2021). Relatedly, it could reduce the space for tinkering, which has been described as a crucial component of research (Knorr, 1977; Knorr-Cetina, 1981), and a critical aspect of 'good research' conceptualised as

good care (Brenninkmeijer & Pols, 2026). The next statement from a PhD at a German neuroimaging lab is related, because while preregistration is often presented as an intervention that makes reporting more specific and explicit, the researcher sees a danger of hypotheses becoming instead vaguer and looser if preregistration becomes the norm:

> Does it mean that then everybody makes very generic predictions? Yes, probably. If you know that you're going to pre-publish something, then you are going to formulate your experimental hypothesis in a more loose, more vague, in a different way than you would do if you were just setting out to do an experiment on your own and nobody needs to know about it (PhD Neuroimaging Lab, Germany).

Time investment and workload for preregistration

As any research practice, preregistration requires the investment of certain resources, among them time and workload. However, as previously mentioned, the amount of time and work necessary to practice preregistration can depend on the type of study that is being preregistered and how detailed the preregistration is going to be. Additionally Open Science practices like preregistration have to be fitted into and arranged with the already existing administrative and bureaucratic tasks most scholars have to handle anyway (Hostler, 2023). Moreover, researchers have explicitly linked the availability of time to the ability of producing 'good science' (Sørensen et al., 2026). Furthermore, if researchers really employ something like an investment strategy (Latour and Woolgar, 1979; Peterson & Panofsky, 2021; Shapin, 1995a), when considering applying a practice or procedure, then it is important to understand how they experience such investment and their expected returns, that come with a procedure like preregistration:

> It takes extra time and extra effort to do a preregistration of your study. I did not preregister my first study, because I was too overwhelmed with everything else (PhD Neuroimaging Lab, Germany).

> It depends a little bit on the complexity of the study of course. So, a behavioural study with just a few outcome measures will be easier to preregister than an fMRI study with a lot of downstream analyses and a lot of potential. There is going to be some bandwidth there, but on average I think I would say between three days and two weeks. It's more than two weeks I would I think it might be too much. Of course, it also depends on how quick the researcher is, but I mean you really think about full-time investment (PI Neuroimaging Lab 1, The Netherlands).

While the first response indicates that preregistration can be perceived as a burden added to an already stressful workload (Hostler, 2023). The second quote suggests the problem that someone can also spend too much time on a

preregistration. Another researcher also states that preregistration and registered reports are 'definitely very cool', but when considering whether to actually practice them the aspects of availability of time and expected workload enter the decision process (Sarafoglou et al., 2022):

> The question, okay do I preregister this study, do I write a registered report for that, which I personally have never done, because it is too much work. But exactly, it is definitely very cool if you do that and if you have the time for that[48] (Postdoc Neuroimaging Lab, Germany).

To get a glimpse into what kind of considerations can go into writing up a preregistration, the following is an excerpt from my fieldnotes during a meeting between a supervisor and PhD student:

> The postdoc mentioned that there is a huge degree of freedom and that they would need some kind of codebook. They want to preregister a strategy and asked whether it would work considering the complexity. The Postdoc looked up a paper in R and they checked in those papers how variables are described that they need for their research. They also looked at the different databases to see and compare how they use/ describe/ code variables. Then they agreed that the PhD student should have something done by Monday. Prior to preregistering the project, they want to contact a colleague with some questions that need to be answered. […] The PhD student asked why the postdoc has to read the literature. The postdoc said needing to know the literature to be able to consult, for instance, on the preregistration. […] The student offers to already write the intro and method section for the preregistration. There are some methodological questions still open that need to be answered: how to use the census? How to use a model? How to link the geographical data? The post doc then says that they might not want to do two new things as compared what exists in the literature. It would, according to the postdoc, be pedagogically more useful to change one thing at a time considering the literature: different method or different model.

Some responses from researchers suggest that the burden of time and work can be addressed by other measures and procedures. Within the neuroimaging institutes the requirement to present a project proposal prior to getting access to the necessary machines is presented as such a measure. One postdoc in a German neuroimaging lab, for instance, describes how it is mandatory in their institute to give a project proposal presentation prior to starting a study to be allowed to use the machines (scanners), which are a hot commodity. In other words, the proposed or planned study is presented to a group of

[48] Original in German: ‚wenn ich mir irgendwie, wenn ich eine Studie plane, die Frage, okay, preregistriere ich diese Studie, schreibe ich dazu ein Registered Report, was ich persönlich noch nie gemacht habe, weil es zu aufwendig ist, aber, genau, ist auf jeden Fall sehr cool, wenn man das macht und die Zeit dafür hat.'

colleagues prior to being able to conduct the study, because it is a requirement to present what is planned in the specific study. In this context, according to the postdoc in their institute the burden to practicing preregistration is smaller, due to a local Open Science Initiative having lobbied and succeeded in making such mandatory project proposals more extensive and something like a 'small preregistration'. The similarity between preregistration and the project proposal meetings at the neuro-imaging labs was also mentioned by a PI from a lab in the Netherlands, calling it 'preregistration-light' (PI Neuroimaging Lab 1, The Netherlands)

> This is actually really cool with this new project proposal. There the Open Science Initiative here at the Institute achieved to make it into something like a small preregistration. So, once you have done the project proposal then you can basically nearly upload it as a preregistration on OSF[49] (Postdoc Neuroimaging Lab, Germany).

Along similar lines a researcher from a cognitive science lab in the UK mentions how making the process of preregistration a part of the normal practice or workflow in the lab can make it more manageable for the researchers involved. The researcher here mentions how closing the distance between the normal or routine work and the work of preregistering decreases the required extra investment for preregistration:

> I think it's more work if you're not used to it. If you kind of standardize it, which I think [PI] has. […] I think it's reducing the barrier to entry […]. Just because it's a good way of doing it is not really enough of a motivation, I think, to get people to do it. I think if people already think what they're doing is good, and the other thing achieves the same thing, basically, they're not going to switch. But if the barrier to entry is low and it's effectively the same, then yeah, sure (PhD Cognitive Science, The UK).

Hence, it is suggested that measures that shape the mandatory or routine part of normal research practice in a way that already prepares researchers for the practice of preregistration reduces the amount of extra investment needed to employ preregistration. By stating that just thinking something is good practice is not enough to change behaviour, but the barrier to entry to the new procedure has to be low enough, this researcher also mentions the crucial distinction, and, at times, confrontation between actual practice of research in the lab and normative arguments about what is good or responsible conduct

[49] Original in German: ‚Das ist eigentlich ganz cool an diesem neuen Projekt Proposal da hat die Open Science Initiative hier am Institut es eigentlich hingekriegt so fast eine kleine preregistration draus zu machen. Also wenn du das Project Proposal gemacht hast dann kannst du es schon quasi so fast als preregistration auf OSF oder so hochladen.'

(practical vs moral). Hence, a certain logic of investment is again hinted toward (Latour & Woolgar, 1979; Shapin, 1995a). However, irrespective of how easy it is to preregister, due to measures that reduce the barrier to entry, preregistration can apparently even be overdone or too detailed and not worth the effort:

> What I have seen sometimes happen is that people would try to preregister every minute detail of the method that they are using. […] Where it becomes complicated is, for example, if you think about fMRI research, sometimes data can be just terribly noisy, for example, because someone fell asleep halfway, or because someone was moving like crazy, or someone was really anxious in the scanner. […], and the list could go on and on and on. And now the problem is if you want to really say beforehand exactly what defines whether you should throw someone out, you have to have that entire list. You have to make choices on all these dimensions beforehand and, I'm not saying it's impossible, but it can be extremely difficult, especially when you haven't done so many fMRI studies. You don't even know all the ways in which a data set can be screwed up. So, you would have to then go around and first become a complete expert on all the possible ways in which your data could be screwed up and so it would take you potentially a year to actually preregister a study. I really think that there the efforts are not paid out by the benefits of the preregistration (PI Neuroimaging Lab 1, The Netherlands).

This further becomes relevant in the context of some proponents of preregistration arguing that it only fulfils its purpose and benefit for the quality and integrity of the research if it is as detailed and comprehensive as possible to be producible and effective (see e.g., see Ofosu & Posner, 2023; van den Akker et al., 2024a). However, here there seems to be the perception that there is a trade-off for researchers between how much time they put into a preregistration to make it as detailed and comprehensive as possible and what they get out of it for their research. Henc, researchers could be able to make preregistration effective, but it might still be perceived as not worth the effort. During the field observation in a conversation with me after a meeting the PI of a neuroimaging lab mentioned having witnessed people wasting half a year on preregistrations and that one should not try to make preregistration into a piece of art. Hence, there seems to be a point of diminishing returns.

Preregistration is maybe not a prison but a punishment

Similar to comments about overly demanding and strict institutional ethical protocols, some researchers also describe their experience with preregistration as burdensome and an annoyance due to experiencing it as inapplicable to their research or feeling more like a prison rather than a plan, due to too strict demands, concerning adherence to a preregistered plan, by peers.

Therefore, here I will shortly elaborate on such cases of problematic experiences with preregistration.

> But I remember just sort of disagreeing [...] with [the] prescriptions, because they made so much sense for memory research. [...]. But what I was trying to do was build a model [...]. So, I was sort of, in an engineering setting, personally. [...] I just couldn't see where on earth [...] preregistration fit in, apart from just the logbooks that I already kept for myself, putting them online. [...] I just don't understand what for me it would be doing (PI Cognitive Science, The UK).

This links to discussions in the literature about how not just the form of Open Science practices varies depending on the type of research and the specific context it is embedded in, but that the relevance and feasibility of such practices is affected by what has been termed epistemic diversity (Bazzolli, 2022; Guttinger, 2020; Leonelli, 2018, 2022, 2023; Malich & Rehmann-Sutter, 2022; Penders et al., 2019, 2020; Ulpts, 2024; Ulpts & Schneider, 2025). Here the researcher talks about that preregistration is inapplicable to the type of model building research or would just be redundant and unnecessary because the researcher describes already keeping track with a logbook (Szollosi et al., 2020). Thus, there seem to be other practices at the disposal of the researcher that already fulfil the function that the researcher ascribes to preregistration and what it could potentially do for their research (no perception of added utility). Previously this aspect of redundancy was put a bit more positively, for example in the context of the project proposal meetings in the neuroimaging institutes, as an aspect that prepares researchers to practice preregistration, but here it is also framed as a form of waste. The scholar concretises that even within the confines of psychology alone the diversity and disunity of modes of knowledge production (Green, 2015; Koch, 1993; Galison & Stump, 1996; Ulpts & Schneider, 2025), leads to a situation where preregistration is seen as inappropriate or redundant for certain psychological research (Flis, 2022).

> Psychology is a very broad church, and there's a canonical version of psychology, which is kind of like an ANOVA. [...] But it bears no resemblance to the kinds of manipulations that I do in my lab (PI Cognitive Science, The UK).

Hence preregistration is seen as one option among many for demonstrating credibility, but it might not be appropriate or the best option considering the context, type and aim of the research in question. This also sheds doubts on existing preregistration incentives and requirements that are applied overly general across the research landscape (see e.g., TOP guidelines: Grant et al., 2025; Nosek et al., 2015). Moreover, it begs for a more nuanced, maybe bottom-up epistemic community specific, enactment of research assessment and evaluation practices without using preregistration as a general or near universal indicator for research quality, credibility, research integrity or good

responsible research conduct, irrespective of the specific epistemic properties of the community and research in question (Hostler, 2024a; Ulpts, 2024).

This leads to a contested topic in the literature concerning the procedure of preregistration and how it is supposed to be used, namely, how to deal with deviations from the preregistered plan (Lakens, 2024). Some proponents of preregistration explicitly proclaim that 'preregistration is a plan, not a prison' (Nosek et al., 2019, p. 817; also see DeHaven, 2017). The question of deviating from preregistrations is closely linked to the confirmatory and exploratory distinction, but not the same. Simply put, how to deal with deviations is part of an overarching question of how to handle the confrontation between a plan and implementation into the often messy and contingent reality that is research (Field et al., 2020; Lynch, 1985). Here I address this by inquiring into how the researchers experience this confrontation. One researcher in the context of preregistration being a component of the master's thesis states that their actual research was very different from what they had preregistered. While actually demonstrating doubt about how to deal with deviations and the influence of constrains presented by limited time and resources available, the scholar mentions that deviating from the preregistration is okay, as long as you transparently report those deviations and indicate how the trajectory of the study changed, because the traceability that results is seen as another benefit of preregistration. However, in the end it is still stated that strictly adhering to the preregistration without deviations is seen as the standard practice of preregistration, indicating some kind of perceived community standard for clean one-to-one implementation of a preregistration, that seem somewhat in conflict to what is doable. This is again in accordance with the often-mentioned assumption of having to know what happens and works beforehand to write a proper preregistration. At the same time, it is suggested that it is an indirect benefit of reporting deviations properly that one can provide a log of what happened.

> The experiment I had presented in my preregistration was very different from the experiment that I had in my master's thesis. Just because things came up and the design changed and doing it in a certain way made more sense as a way of answering the question that I was originally trying to answer. […] there are still limitations to the amount of time you have, the amount of resources you have. […] But the preregistration is there, so you can see the progress of how it changed. […] It's kind of like a log of the way things progressed. And I guess that's like an indirect benefit of doing it that way. Even if you don't follow the standard practice of preregistration and committing to it (PhD Cognitive Science, The UK).

Next a researcher shares a somewhat problematic and expensive early experience with preregistration, in the context of their responsibility as a PhD

supervisor. The scholar outlines how a former PhD student had preregistered certain, apparently too demanding, exclusion criteria. This, while trying to adhere to the preregistration, lead to having to run the study longer than intended and spending a lot of money on getting enough participants. The problems arose, according to the researcher, because the preregistration was written with 'a pretty bad idea' of what to expect during the actual study. Here the researcher also describes feeling somewhat 'stuck' or 'trapped' in a badly specified or inappropriate preregistration that the authors set out to follow. What follows is quite the long quote, but it nicely summarises such an experience of feeling imprisoned and punished by a preregistration.

> You can fall into a trap of trying to write things down before you do them, but you don't yet have any idea [...], and then you write down something that's a pretty bad idea. And then you later feel like you obliged yourself to do something that you shortly after realized was a bad idea. And then you ended up stuck with it for ages. [...] I was co-supervising a PhD student [...] [PhD] preregistered a set of exclusion criteria for the studies that, had I looked at them, I'd be like, they're very stringent. [...]. And then we started running these studies which were very expensive, and we did stick to our guns. We honoured these exclusion criteria, but they led to a 65%-70% rejection rate of all our participants. We spent thousands and thousands of dollars on participants. We threw away who I would say most of which were actually attending and giving us good data. [...] I felt hamstrung by a poor decision that had been written down by somebody who hadn't thought about this yet, because they hadn't even tried it (PI Cognitive Science, The UK).

The PI goes on to somewhat describe this bad early experience as a learning experience where afterwards the scholar recommends to only preregister research where one has a clear idea of how something is going to come out and what your hypotheses are. According to the scholar one should not feel obliged or pressured to fill in the preregistration form and just make up hypotheses or other details when one does actually not have a clear notion of what is going to happen. Lastly the cognitive science PI also mentions that this could actually make the 'scientific process worse', with preregistration being in danger of becoming the metrics and distracting from the real problem (see e.g., Klonsky, 2025).

> The better way to understand OSF is, if you don't have a hypothesis, don't just make one up on the spot to fill in the form, because all you're doing is making your scientific process worse. [...] There's this danger of the metric overtaking the real problem (PI Cognitive Science, The UK).

A personality psychology professor in Germany also identifies certain problems with preregistration by providing their own personal experiences. This

researcher, for instance, outlines being confronted with preregistration during peer review in a way that can be experienced as burdensome or 'annoying' and how preregistration as a practice does not necessarily agree with the professor's conception of good or credible scientific practice. Moreover, here again the aspect of preregistration being seen as redundant in the context of already existing practices comes up. Specifically, the professor deems robustness demonstrations enabled by the sharing of data as more valuable and credibility inducing than following one preregistered plan, because then everyone can run their own analysis.

> On my account write down the pipeline, then say where you deviate from it. However, […] I find that nearly a bit annoying. Especially when the reviewers then request, but please every single thing detailed and in the text and first the results for the preregistered analysis. My counterargument is, if I share the data, report the best possible analysis, then everyone is able to take the data and run their analysis that is possibly better, and also find out how robust it is. So my idea would rather be to do robustness analyses then. […] I notice for myself when I am convinced from a finding or not convinced, then I try to get that finding in different ways[50] (Professor Personality Psychology, Germany).

The researcher also seems to refer to a conflict that exists between what some proponents see the ideal usage of preregistration to be and how it is applied in practice. Even if it is often claimed that deviations are in principle allowed and not viewed as problematic or indicating research of a lesser value, in practice reviewers and researchers seem to apply preregistration differently and use deviations from preregistrations as an indicator for lower quality, integrity or value of the research. However, then it does not seem to matter too much what proponents of preregistration see as the ideal use. In other words, while some proponents of preregistration clearly state that deviations are okay, as long as you transparently communicate them (see e.g., Lakens, 2024, p. 1 for 'when and how to deviate from preregistration'), the experience of researchers with

[50] Original in German: ‚Mein persönliches Ding ist immer noch, naja, meinetwegen schreibt man halt dann die Pipeline. Dann sagt man die Abweichungen davon. Ich finde es aber manchmal fast schon ein bisschen nervig. Ja, das ist eine Lösung. [Postdoc] macht das auch so ein bisschen so. Aber ich finde das fast schon ein bisschen nervig. Besonders wenn die Reviewer dann kommen. Aber bitte jedes einzelne ganz genau und auch in dem Text und erstmal die Ergebnisse für die Präregistrierte-Analyse. Mein Gegenargument ist, wenn ich die Daten zur Verfügung stelle, die bestmögliche Analyse berichte, hat jeder die Möglichkeit, die Daten zu nehmen und seine Analyse, die möglicherweise besser ist, drüber zu rechnen. Und auch rauszufinden, wie robust das ist. Also meine Idee wäre eher, Robustheitsanalysen dann zu machen. So gehe ich auch ran. Also ich merke selbst, wenn ich einen Befund überzeugt bin oder nicht überzeugt bin, dann versuche ich ihn auf verschiedenen Wegen auch zu erreichen.‘

preregistration seems sometimes different, because such an ideal is not represented in the reality of scientific practice in which researchers have to navigate in the job marked and practice their research (Elizondo & Kaltenbrunner, 2024). Furthermore, one might ask whether the extra amount of work that the researcher refers to resulting from the 'annoying' demands by other scientists to report the preregistered part and then in detail explicate every deviation, can actually feel like punishment. Hence, while it might be stated that 'preregistration is not a prison, but a plan' (Nosek et al., 2019, p. 817), it can still feel like deviations from the preregistration are actually punished in practice, because such deviations result in demands for explanation and sometimes extensive extra work or even result in devaluation. Thus, preregistration does not have to be a prison, but it can still feel like a ball and chain you have to drag with you.

Another issue that surfaces with preregistration is that it is difficult or even impossible to (properly) preregister complex or complicated research. This was previously already hinted at when researchers mentioned that the workload related to preregistration can vary depending on the complexity of the data and the research, while there also seems to be a point where the amount of work does not pay off. This is mentioned as especially troubling when complete adherence to the preregistration is demanded:

> There are proponents (followers) of preregistration who think for it to work you have to provide the exact analysis-pipeline beforehand. On the one hand I can understand that, but on the other hand, I think that doesn't work. Because some things I only realize by analysing. My analysis is getting better by analysing the data. Because only then I recognize, for example, the distribution is different. Or I have outliers here according to criteria that I did not state a priori[51] (Professor Personality Psychology, Germany).

This response suggests somewhat of a dilemma in the researchers view on preregistration. On the one hand, there is an understanding for demanding a

[51] Original in German: ‚Meine Erfahrung is aber, dass das, wenn die Analysen komplex sind, wenn die Daten schwierig sind, wenn man Paneldaten bekommt, schwierige Daten, auch gerade schwierige statistische Analysen, dass man immer irgendwas nicht genau richtig preregistriert. Und dass es Anhänger gibt, der Preregistrierung, damit die eigentlich richtig funktioniert, muss man genau die Analyse-Pipeline vorhergeben. Das kann ich auf einer Seite verstehen und auf der anderen Seite, glaube ich, funktioniert das nicht. Weil manche Sachen mir erst durch die Analyse klar werden. Meine Analyse wird besser, indem ich die Daten analysiere. Weil ich dann erst merke, die Verteilung ist beispielsweise anders. Oder ich habe hier Ausreißer nach Kriterien, die ich vielleicht nicht a priori festgelegt habe. Gerade sowas wie Verteilungssachen. Werden mir manche bestimmte Sachen dann erst klar oder es gibt ein Model-Misfit. Darauf muss ich reagieren oder ich sollte rausfinden, an was es liegt.'

comprehensive preregistration prior to the analysis, but on the other hand it is seen as unrealistic since unexpected things will inevitably happen. This is further the case because analysis is also described as a learning process, where the research improves on the go while conducting the analysis. This seems like statements made by researchers in Pols and colleagues' (2024) study about what constitutes 'good science', where 'good science' is perceived as a collective learning endeavour and characterised by a need for space for the unexpected. Therefore, that researcher goes on to say that preregistration might make the research with complex analysis worse but also sees the danger of p-hacking. Then this is described as a trade-off in the practice of preregistration between having the freedom to adapt and explore analysis and limiting that freedom with preregistration because of a danger of exploitation:

> That means I just think stating the exact analysis beforehand when the analyses are complex worsens the research. That is my fear. However, I also see that if I don't adhere to it, I do, of course, have the danger of p-hacking. […] Yes, that is a trade-off somehow[52] (Professor Personality Psychology, Germany).

In line with earlier sentiments about 'good research' and OS not being the same and there being space for conflict between the two (Chapter 5), crucial seems the subsequent sentiment that the researcher utters as standing alone against Open Science, because it suggests that the professor already experiences something like a culture shift, which is perceived as, at least somewhat, problematic or suboptimal in the context of which practices actually enhance or communicate credibility according to the researcher. Touching again on the circumstance that research is not always confirmatory, and preregistration seems to be perceived as contributing to making research more orthodox or generic. While the scholar states the wish to also do something novel and again voicing annoyance of being burdened during the peer review process with having to painstakingly explicate every deviation. Importantly, this annoying experience is seen as somewhat inevitable, because it is just 'logical according to the preregistration logic'.

> But I also want to have new ideas, I also want to find something new, I want to think about something new. I don't know whether that is in the end, that is all against Open Science. […] Whether this will really be the best way in the end, I am not sure. But I think I am relatively alone with this view. […] I still preregister because it also doesn't do harm as long as I am allowed to deviate from it and the

[52] Original in German: ‚Das heißt, ich glaube, einfach bei komplexen Analysen vorher genau die Analyse darzustellen, verschlechtert die Forschung. Das ist meine Befürchtung. Aber ich sehe auch, dass wenn ich mich nicht daran halte, ich natürlich die Gefahr von P-Hacking habe. Wenn es so ist, dass ich verschiedene Analysen ausprobiere und am Ende das veröffentliche, das rauskommt. Ja. Das ist irgendwie ein Trade-Off.'

> reviewers don't demand that I first write 3 pages in the article stating what I said I would do, then I write a paragraph, oh by the way, everything was wrong as of now, and now the correct stuff follows. However, I have read such things already, once I was even present when exactly that happened. However, this is logical according to the preregistration logic[53] (Professor Personality Psychology, Germany).

Overwhelming students with preregistration

Another perceived issue with preregistration that comes up and is related to being, at least to a certain degree, able to anticipate what will happen during the actual study and how everything works, is preregistration for students. As the experience of the PI, from the cognitive science lab, who felt 'hamstrung' by a PhD's bad or too ambitious preregistration suggested, preregistration by or even demands for students to preregister seems to be seen as largely problematic, because it might be unrealistic to expect from them to be able to have enough experiences with specific procedures or analyses techniques. This apparent lack of knowledge that might be necessary to do a proper preregistration can, next to the danger of feeling chained by a faulty preregistration, according to the respondents, lead to a situation where the people involved become overburdened.

> It is really difficult for students, [...] they are supposed to preregister everything in social psychology. They are completely overwhelmed, because they are still in their education. They once had statistics 2, never calculated something like that and are now supposed to decide beforehand in their heads what to calculate. This is also didactically completely unfortunate. And there is a working group where everything has to be preregistered. But then I would say again, that undermines preregistration. If I have 10 bachelor's theses a year and every two years something interesting comes out of it and all 20 are preregistered and I then

[53] Original in German: ‚Aber wenn es komplexer ist und ich gar nicht genau weiß und Forschung ist nicht immer nur konfirmatorisch und ich prüfe nur und mache genau das gleiche, sondern ich will ja neue Ideen auch haben, ich will ja was Neues finden, ich will was Neues überlegen. Weiß ich nicht, ob es am Ende die... Das ist alles gegen Open Science. Jetzt alles anders. Ob das wirklich am Ende der Königsweg sein wird, bin ich nicht sicher, aber ich glaube, ich stehe damit relativ alleine. Was trotzdem dazu führt, alle diese Sachen, irgendwie präregistriere ich dann doch, weil es schadet auch nicht. Solange ich mich erlaubt bin, davon abzuweichen und solange die Reviewer nicht verlangen, dass ich erst einmal drei Seiten im Artikel schreibe, nämlich genau so machen, wie ich es gemacht habe, dann schreibe ich einen Absatz, ach übrigens, und alles war falsch, wie ich heute war, und jetzt kommt mal das Richtige. Aber ich habe so Sachen schon gelesen, einmal war ich sogar dabei, da ist genau das passiert. Ich finde das schlecht für den Artikel. Aber es ist logisch, in der Präregistrierungslogik muss es so sein. Also vielleicht passt auch besser wirklich so in experimenteller Forschung, aber auch da kann was passieren. Was man vorher nicht wusste'.

> publish it as preregistered, that didn't help me at all. Since that can as likely be a p=0.05 chance finding. That's why the demand that every bachelor thesis has to be preregistered is also again such an Open Science demand, in practice it has to be supervised very well[54] (Professor Personality Psychology, Germany)

In the end the scholar seems to describe another form of PARKing where preregistrations of bachelor's theses can be used as a fishing trip that will turn out a chance finding at some point that then will be published as preregistered (Klonsky, 2025; McDermott, 2022; Yamada, 2018). This could also be seen as a form of p-hacking aided by preregistration instead of prevented by it. We could also say that this is a way to exploit students to get a clean publication for the lab. Here the researcher further hints to an extra amount of workload for the supervisor in cases where preregistration is demanded from students, because there seems to be a need for a monitoring body for preregistrations and need for supervision to compensate for the lack of expertise and experience so that students can preregister productively and effectively. Such monitoring is further seen as required to ensure that such procedures as well as students are not exploited. Next a postdoc from a German neuroimaging lab also talked about how it was 'extremely difficult' as a student to preregister, because there was not enough experience to write the preregistration in proper detail:

> Back then I found it extremely difficult before my first study, because it wasn't as clear to me back then, and even afterwards, to be honest[55] (Postdoc Neuroimaging Lab, Germany).

---

[54] Original in German: ‚Und dann habe ich festgestellt, bei Studierenden, das ist dann ganz schwierig, wenn die jetzt kommen, und dann kommen die gerade von mir beraten, weil ich statistisch ein bisschen was kann, sollen der Sozialpsychologie alles präregistriert werden, sind völlig überfordert damit. Weil noch in der Ausbildung, die haben mal Statistik 2 gehabt, haben noch nie sowas gerechnet und sollen jetzt vorher, im Kopf entscheiden, was sie wohl rechnen werden. Das ist auch didaktisch völlig unglücklich. Völlig überfordert. Und da gibt es eine Arbeitsgruppe, da muss alles präregistriert werden. Dann würde ich aber auch wieder sagen, dann wird eine Präregistrierung unterlaufen. Wenn ich nämlich 10 Bachelorarbeiten pro Jahr habe und jedes zweite Jahr kommt was Interessantes raus und alle 20 sind präregistriert. Und ich veröffentliche das dann als präregistriert. Da hat es mir überhaupt nichts geholfen, weil das genauso p=0.05 Zufallsfund Also, deswegen auch diese Forderung, jede Bachelorarbeit muss präregistriert sein, wie, also es ist auch wieder so eine Open Science Forderung. In der Praxis, es muss sehr, sehr gut betreut sein dann auch.'

[55] Original in German: ‚Ich fand das damals vor meiner ersten Studie extrem schwierig, weil mir das alles nicht so klar war und auch nachher war es ehrlich gesagt so also dann waren die Daten da und wir haben uns quasi nochmal neu gefragt, ja und wie analysieren wir das jetzt? Und es war nicht so glasklar wie es sein sollte glaube Ich.'

The complication of having students preregister or even write registered reports also came up in a conversation during my field observation with the PI of the German neuroimaging lab. The PI discusses how registered reports are most appropriate for established researchers who are already immersed in a topic and have experience, because for PhD students it can be a bit much and overwhelming to have to write up a whole theory section (introduction) and method section without having done the research. Furthermore, the sentiment, that such procedures are most applicable for replication studies reemerges. During the conversation the PI also mentions attending a workshop where people talked about their experience with registered reports and one of the conclusions was that it seemed very appropriate and useful for replication studies, when the methods and theory can to a large degree be taken over from the study that is supposed to be replicated.

Preregistration: a story of credibility and hardship

As the field observations and interviews reveal the practice of preregistration seems to be perceived as an intricate combination of hardship, due to additional workload, pressure to conform to a pre-made plan, and the potential inapplicability as well as gameability, with the utility of functioning as a form of reporting that aids in the management and communication of credibility by upholding the distinction between exploratory and confirmatory components of the research and constraining the possibility for the exploitation of researcher degrees of freedom, for instance by controlling bias, in addition to helping in planning out a study from the beginning.

The researchers' responses suggest that the perceived benefit of preregistration largely depends on it enabling them to put oneself out there by sharing one's plans concerning the steps and order of the research so that others as well as oneself can check and control that everything is going according to plan and that deviations are accounted for. Hence, while actual experiences of being checked by others through their preregistration is not really described by the researchers or in a rather negative way, this aspect of preregistration of providing the potential for researchers to check or control themselves and potentially being checked by others, seems to be the underlying reason for the perceived advantages of preregistration. Only the aspect of supporting researchers in producing a well-thought-out study seems independent of this underlying motivation of surveillance and accountability, but for that purpose they would not need to put the preregistration online for everyone to see and potentially control, something like a logbook would do just fine. However, even for that you seem to require time as well as a certain topic and technology specific (tacit) knowledge and experience to be able to achieve that goal. To some degree preregistration is described by some as a procedural or

bureaucratic reporting tool to ensure that human bias and temptation to intervene does not contaminate the data or the scientific process, by exploitation of degrees of freedom, and thereby protect its integrity. One could argue that this thriving for controlling or even preventing human bias and temptation to intervene in the scientific processes by making researchers follow a pre-made plan and allowing the community to hold one accountable is part of a moral economy of mechanical or noninterventionist objectivity (Daston, 1995; Daston & Galison, 1992; Field & Derksen, 2021; Penders, 2022). In this context, one can demonstrate to the community one is part of that one is credible and a good scholar by conveying that there was no tempering with the data or scientific process at large. Whether induced by the Open Science reform or not and it was already characteristic of this epistemic community before, preregistration seems for most to be part of a conceptualisation of research integrity and 'good science' according to a moral economy of mechanical objectivity (see e.g., Penders, 2022).

Within what Latour and Woolgar (1979) describe as the cycle of credit, credibility is one necessary type of capital that is converted into other types of capital, like funding, to survive in the academic market. In Shapin's words preregistration is described as a credibility management strategy and a literary technology that aims to make the rest of the research community virtual witnesses and thereby functions as a procedure that can persuade them of the research's and researcher's credibility to gain said credit paid for through credibility (Latour & Woolgar, 1979; Shapin, 1984, 1995a; Shapin & Schaffer, 1985). Meskus et al. (2018), argue that OS speeds up the multiplication of witnesses (Shapin, 1984; Shapin & Schaffer, 1985) and in the context of the STAP Stem Cells misconduct case, call the appraisal in the associated online discourse, in blogs, social media, and post-publication peer review fora, 'accelerated virtual witnessing' (Meskus et al., 2018, p. 5). Accordingly, the practice and appraisal of preregistration can be conceptualised as an act of credibility work that help in the social process of persuading the research community that they and their research are credible (see Chapter 3).

This seems to be happening within a culture of suspicion in contrast to a culture of trust (also see Pulverer (2014) for warnings against a vigilante culture in science), because the research is seen overly sensitive to human interventions and biases. In that context it aids in the social process of persuading the community that the research and the reported findings are credible, by demonstrating a willingness to exposing yourself to that suspicion. On a more critical note, in reference to Schniedermann (2025) one could argue that the spread of preregistration is part of the standardisation of scientific (biomedical) reporting practices with it being a bureaucratic or procedural patch over the task uncertainty inherent to psychology and adjacent disciplines in an

image of ‘post human’ science (Peterson & Panofsky, 2023, p. 164; Penders, 2022; Whitley, 1984). As Klonsky (2025) and other caution this can lead to a situation where preregistration itself becomes to be used as an indicator for ‘good science’ or research integrity, potentially leading to a research reality where preregistration functions not as a credibility enhancing practice but as a superficial veneer of credibility (Devezer et al., 2021; Devezer & Penders, 2023).

The responses indicate that the researchers do indeed follow, or retrospectively justify with, an investment strategy when considering to preregister or use the registered reports format in their research with a specific trade-off between the amount of resources like time and workload needed being weighed against the potential benefit in the sense of credibility and advantages in the academic job market (Davies, 2019; Hostler, 2023; Latour & Woolgar, 1979; Peterson & Panofsky, 2021; Shapin, 1995a). While required workload and time investment for preregistration seem to be reducible by packaging the necessary work or making preregistration part of the day-to-day workflow to increase feasibility, there are still issues of relevance due to epistemic diversity and different conceptions of what is ‘good’ or credible science (see e.g., Szollosi et al., 2020; Ulpts & Schneider, 2025). Interestingly, some of the responses, especially in the context of piloting, suggest that preregistration instead of reducing or constraining researcher degrees of freedom might actually provide the playground for even more degrees of freedom. Furthermore, as the ethnographic work in end-of-life research from van Drimmelen and colleagues (2024) suggests, whether protocols are underdetermined or overdetermined researchers’ discretion remains inevitable, because adaptations are necessary for local contexts and additional decisions have to be operationalised once research is executed. Thus, degrees of freedom are moved somewhere else, but not eliminated.

The researchers further indicate that having students preregister their work might actually overburden them instead of helping them. Moreover, we might also ask whether the potential exploitation of students in the context of preregistration is a symptom of an underlying notion of mechanical objectivity. As Daston put it: ‘It is this craving for authenticity that also explains why some scientists toyed with the idea of employing uneducated assistants. Not only was their labor cheap; they were, it was thought, nearly as much tabula rasa as machines were, and therefore more fit to record observations than the all- too-well-informed and expectant scientist’ (Daston, 1995, p. 20). Even though it is claimed by proponents of preregistration that it is to be used as a ‘plan, not a prison’ (Nosek et al., 2019, p. 817), in practice some of the researchers report experiences that they felt stuck with bad decisions that were set in the preregistration protocol or burdened by the amount of extra work

that resulted from demands to explicate and explain every deviation from the preregistration. This was explained by a perceived demand or norm within the community for adhering to the preregistered plan as closely as possible. However, this begs the question of how relevant guidelines and fine print in recommendations about the nuances of preregistration are, when the practiced realities and perceived demands and norms differ so much. Hence, while not necessarily a prison, preregistration apparently still can feel like a ball and chain.

All in all, preregistration is described as an intricate combination of helping in demonstrating credibility and experiencing hardship. Additionally, further research and clarity seem to be required when or whether the necessary investment for a proper preregistration is worth the effort or even relevant according to actual researchers in different epistemic communities. It, for instance, remains questionable how useful preregistration is perceived in research cultures that are less build on suspicion, but more on trust, like some qualitative research approaches (Moors, 2026; Pachirat, 2018). Lastly, policymakers, reformers and research integrity researchers alike must be careful not to let preregistration become another thoughtless or ritualistic indicator of research quality or research integrity, preventing actual assessment of credibility based on the quality of the research (Devezer et al., 2021; Devezer & Penders, 2023; Klonsky, 2025). Like mindless statistics (Gigerenzer, 2004), or ongoing discussions about problematic metric-based research evaluation (Rushforth & Hammarfelt, 2023) and Campbell's law (Campbell, 1979), such mindless use of preregistration might actually worsen the state of science instead of improving it.

Chapter 8: The practices of replication and reproducibility in the lab

In this chapter I provide an understanding of research integrity as 'good research' through the researchers' stories of how they engage with practices that have been linked to research integrity as instruments for credibility work and management within their research community. The practices in question, are reproducibility and replication. I first give a short introduction about the literature on the relationship between replication, reproducibility and research integrity. Next, I will outline sociology of scientific knowledge (SSK) work on the practice and appraisal of replication. Afterwards, in the empirical analysis I will explore the enactment of replication and reproducibility in the lab through the stories that scholars tell me. In the context of a prevailing conceptual confusion surrounding the replication and reproducibility terminology, I first delineate the scholars' usage and understanding of these concepts followed by the epistemic functions they associate with them (Gläser, 2024; Shapin & Schaffer, 1985; Ulpts & Schneider, 2025). Subsequently, I present the researchers' stories of their own replication and reproducibility practices as well as their experiences with failed replications and the perceived dilemma between incentives for novelty and practicing replications. To conclude I will relate the findings to sociology of scientific knowledge research on replication and reproducibility to comprehend the role of replication and reproducibility in the management of credibility.

Introduction

Most likely due to the increased attention the topics of replication and reproducibility have received in the context of recent claims about a 'replication' or 'reproducibility crisis' (see e.g., Baker, 2016; Peterson & Panofsky, 2023; Romero, 2019; Wiggins & Christopherson, 2019), and calls for increased replication and reproducibility as one of the 'hallmarks of good science' (Resnik & Shamoo, 2017, p. 116) or the 'scientific gold standard' (Jasny et al., 2011, p. 1225) by prominent metascience and Open Science proponents alike (see e.g., Buck, 2024; Edlund et al., 2022; Haig, 2022; Ioannidis, 2012; Nosek et al., 2012; Rubin, 2023; Simons, 2014; Zwaan et al., 2018), in the current research integrity literature reproducibility and replication are among the discussed and promoted practices and standards aimed at securing as well as improving research integrity (see e.g., Bouter & de Riet, 2021; Diaba-Nuhoho & Ampon-

sah-Offeh, 2021; MacLeod, 2022; NASEM, 2017; Peels, 2019; Peels & Bouter, 2018; Resnik & Shamoo, 2017; Ross-Hellauer et al., 2022; Tang, 2025a; Tang, 2025b). Accordingly, within the research integrity discourse reproducibility and replication are often practically treated like demarcation criteria and diagnostic tools to assess credibility (see e.g., Bouter & de Riet, 2021; Diaba-Nuhoho & Amponsah-Offeh, 2021; MacLeod, 2022; NASEM, 2017; Nosek et al., 2012; Peels & Bouter, 2023; Resnik & Shamoo, 2017; Tang, 2025a). Furthermore, as part of Open Science implementations into responsible research(er) assessment reform, COARRA (2025) does refer to the importance of replication and reproducibility for research quality. Moreover, many policy documents and guidelines for proper conduct around the globe already incorporate OS and reproducibility narratives (see e.g., ALLEA, 2023; Labib et al., 2025; DFG, 2025; Laine, 2018; KNAW, 2018; UKRIO, 2025), a recent example is US research policy which took up the Open Science story, declaring that *Gold Standard Science* is reproducible (The White House, 2025). Maybe not as much in the foreground, but even prior to the influence of the modern OS movement and metascience has replication been a crucial component of the research integrity literature and policy. In the *Introduction to the Responsible Conduct of Research* from ORI, Steneck (2007) states that reporting of research should enable others to replicate the research, by providing sufficient detail. Even earlier, the 1998 memorandum on *Safeguarding Good Scientific Practice* from the DFG (1998/2026), in reaction to the Friedhelm Herrmann misconduct case in Germany, also included replication and reproducibility as crucial components of research integrity assurance.

Similar to a lot of OS reformers and metascience proponents (see e.g., Derksen, 2019; Earp & Trafimow, 2015; Haig, 2022; Nosek et al., 2012; Rubin, 2023), Resnik and Shamoo (2017), in reference to Popper, call reproducibility, as an indicator that findings are objective and reliable instead of due to chance or bias, 'one of the hallmarks of good science (Popper 1959)' (p. 116). They, argue for the importance of reproducibility for research integrity, by recounting the role reproducibility played in misconduct cases. They, for example, mention the 'Baltimore affair', in which O'Toole accused Imanishi-Kari of data fabrication and falsification after being unable to reproduce some of the experiments herself and finding inconsistencies between lab notes and the published record (for a somewhat more extensive perspective on the Baltimore affair and what role scientific hierarchies, anti-science narratives, denial, and bureaucracy played, see Raven & Marks, 2009). While first being found guilty in 1994 by the Office of Research Integrity, in an appeal Imanishi-Kari was later exonerated by the U.S. Department of Health and Human Services and the Office of Research Integrity was criticised for their handling of the case as well as their inability to prove misconduct (Kaiser & Marshall, 1996). Another

one of their examples is the stem cell misconduct case of Obokata in 2014 (Resnik & Shamoo, 2017). Colleagues at the RIKEN Center, where she worked, were unable to reproduce her findings, leading to accusations of misconduct that lead the RIKEN Center to investigate the case and subsequently finding her guilty (Cyranoski, 2014; Resnik & Shamoo, 2017). However, what Resnik and Shamoo (2017) fail to acknowledge, is that suspicion did also arise, because their claims about triggering pluripotency appeared extremely extraordinary to the rest of the stem cell community, as well as the heavy incorporation of the online community in detecting and handling this misconduct case in an era of Open Science (see e.g., Meskus et al., 2018).

Peels and Bouter (2023), acknowledging the diversity in conceptualisations and uses of terminology across fields and over time (see e.g., Barba, 2018; Matarese, 2022; Vachon et al., 2021), link reproducibility and replication to the trustworthiness of findings within a specific scientific community. They define reproduction as the re-analysis of the data from the original study based on the same research protocol. This checks for, what they call, 'numerical consistency', while the variation of the analysis protocol would check for robustness (Peels & Bouter, 2023, p. 79). Then a direct replication is the analysis of newly collected data based on the same analysis protocol as the original study aiming to rule out chance findings, while conceptual replication is the analysis of newly collected data based on a slightly different analysis plan, aiming to answer the same research question and ruling out bias (Peels & Bouter, 2023; also for conceptual work on these terms see Derksen & Morawski, 2022; Hudson, 2023; Schmidt, 2009). Moreover, they state that the notion of successful replications is ambiguous, but also that successful replication increases the trust in the original study's primary findings (also see Nosek & Errington, 2017, 2020; Sarisoy, 2025). All in all, they 'wholeheartedly embrace' that replication is more and more becoming a central topic in discussions about proper science (Peels & Bouter, 2023, p. 85).

Similar to the diversity and confusion concerning what constitutes QRPs, research misconduct, or 'good research' (see e.g., Dal-Ré & Marušic, 2025; Field et al., 2024; Horbach & Halffman, 2017; Kolstoe & Pugh, 2024; Penders et al., 2009; Pols et al., 2024; Ravn & Sørensen, 2021), there is not just a certain diversity regarding which forms replication and reproducibility can take depending on the specific research context (Cole et al., 2024; Leonelli, 2018, 2023; Penders et al., 2019; Ulpts & Schneider, 2025; Ulpts et al., 2025), but also more generally a conceptual confusion, because in the current reproducibility and replication discourse terms and uses seem to proliferate while it often remains unclear what exactly these terms refer to (Barba, 2018; Ulpts & Schneider, 2024). There is a vast plurality and no agreement concerning, on the one hand, what are the practices involved, for instance, whether something

is supposed to be redone, which parts of previous research are supposed to be redone and how similar these components have to be compared to the original study, and on the other, what epistemic functions these practices are supposed to fulfil, like checking or establishing calibration, generalisability, reliability, and validity (see e.g., Albertoni et al., 2023; Barba, 2018; Cole et al., 2024; Fletcher, 2021; Gundersen,2021; Matarese, 2022, 2023; Orkeny, 2025; Sarisoy, 2025; Ulpts & Schneider, 2024).

Science reformers and research integrity scholars often acknowledge that a replication study will always differ from the original study in some ways and there are ongoing disputes concerning what counts as a successful replication (Nosek & Errington, 2017, 2020; Peels & Bouter, 2023; Sarisoy, 2025), some also go so far as to argue to outright abolish the direct-conceptual distinction (see e.g., Feest, 2019; Nosek & Errington, 2020). However, while these reflections on the complications or even impossibility of direct and exact replication in actual practice exist, especially in the biomedical and social sciences (see e.g., Brenninkmeijer et al., 2025; Collins, 1985; Derksen & Morawski, 2022; Feest, 2019; Nosek & Errington, 2020; Stroebe & Strack, 2014), the current science reform and research integrity literature still shows a high focus on direct replications as validation and diagnostic procedures as part of science's self-correction capacity (see e.g., Buck, 2024; Macleod, 2022; Makel & Plucker, 2014; Pashler & Harris, 2012; Simons, 2014; Zwaan et al., 2018; Peels & Bouter, 2023; Peterson & Panofsky, 2021; Resnik & Shamoo, 2017). This focus on direct over other forms of replication has also made it into science governance, as the NWO Replication Studies funding program demonstrates (see e.g., Derksen et al., 2024). Recent endeavours aimed at improving the replicability and, thereby, credibility of research are largely based on processes of explication and procedures of sharing that are supposed to enable others to trace, reuse and ideally replicate the research (see e.g., Announcement: Reducing our irreproducibility, 2013; Nosek et al., 2015; Penders, 2022; Steneck, 2007). Some also actively investigate whether expertise and the involvement of the original researchers are necessary ingredients for getting replications to work, largely concluding that, in the investigated cases, they are not (Ebersole et al., 2020; Klein et al., 2022; Protzko & Schooler, 2020; Ulpts et al., 2025).

Penders and colleagues (2020) conceptualise the recent (re)focus in Open Science (and research integrity) on the role and virtuousness of replication as part of a 'moral economy of repair' (p. 107) (Daston, 1995), where replication is presented as one or even the prime procedure for the diagnosis of error, statistical flukes and misconduct, ultimately leading to the correction of the scientific record (also see Ankel-Peters et al., 2023; Flis, 2022; Haig, 2022; Peterson & Panofsky, 2021; Resnik & Shamoo, 2017; Rubin, 2023), all in the

context of assumptions about a broken science system and a narrative of science in crisis (see e.g., Bartscherer, 2026; Bartscherer & Reinhart, 2025; Peterson & Panofsky, 2023).

However, much of current discourse about replication and reproducibility sits in stark contrast to science studies work, in the tradition of the sociology of scientific knowledge (SSK), on the practice and appraisal of replication and reproducibility. The wider science studies literature indicates that reproducibility and replication are plagued with various practical, epistemic, social, normative and conceptual issues, making it a more complex component of the diverse scientific enterprise than the research integrity literature tends to make it seem (see e.g., Ashmore, 1988; Brenninkmeijer et al., 2025; Chen, 1994; Collins, 1974, 1975, 1985, 1991, 2016; Feest, 2016, 2019; Guttinger, 2018, 2019, 2020; Leonelli, 2018, 2022, 2023; Mulkay, 1984; Mulkay & Gilbert, 1986; Norton, 2015; Penders et al., 2019, 2020; Peterson & Panofsky, 2021; Reinhart, 2016; Shapin & Schaffer, 1985; Travis, 1981; Ulpts et al., 2025). The interview study from Hangel and Schickore (2017), for instance, illustrates just some of the issues that emerge in the context of practicing replication and reproducibility as well as how it is clouded in uncertainty requiring the researchers' discretion, those are (Schickore & Hangel, 2019): whether to use calibration or replication when confronted with new and surprising findings, how to proceed when failing to replicate, how many repetitions are required to trust a finding, how to assess the outcome of a replication study, it can be extremely expensive to do replications which is often coupled with a lack of (perceived) pay-off for replication (see e.g., Peterson & Panofsky, 2021; Ulpts & Schneider, 2025), they are perceived as boring (Mulkay & Gilbert, 1986), no time to ensure or check reproducibility and determine whether something works (Hostler, 2023), unclear how much differences are permissible for something to still count as a replication, replications are seen as hard to publish, the link between replicability or reproducibility and error or fraud is unclear, and replications as an attempt at fraud detection are perceived as non-collegial (Hangel & Schickore, 2017; Schickore & Hangel, 2019). Such science studies attempts at emic perspectives might further help in understanding why, some claim that reforms are bringing a credibility revolution (see e.g., Korbmacher et al., 2023; Vazire, 2018), while others assert that the reproducibility reforms are not working (see e.g., Nelson, 2021; Peterson, 2021). Therefore, to contextualise and further illustrate where and how the research integrity community might be able to use insights from science studies, like the sociology of science (sociology of scientific knowledge) that aim to provide an emic understanding of replication and reproducibility, I will first provide some background and a brief elaboration on existing science studies (SSK)

work on the role and position of practicing replication and reproducibility in science.

Science studies on the practices of reproducibility and replication

The probably best-known contribution in SSK on the practice and role of replication stems from Harry Collins. Although Collins famously called replicability the 'supreme court of the science system' (Collins, 1985, p. 19), he emphasises the complex social, personal, and uncertain nature of replication in practice. Collins' most influential works (see e.g., 1974, 1975, 1985, 1991, 2016) are on the experimenters' regress, the reliance on tacit knowledge, and the role of social negotiation in replication practice and appraisal (also see Brenninkmeijer et al., 2025; Feest, 2016; Shapin & Schaffer, 1985; Travis, 1981). He mostly based his work on his immersion into the (early) gravitational wave physics community. Collins (1975) distinguishes between the algorithmical and the enculturational model of understanding replication. According to the algorithmical model, it is possible to explicate and share all the necessary and sufficient information that can be used like instructions and enable scholars to replicate someone else's research (Brenninkmeijer et al., 2025; Collins, 1975; Travis, 1981). This seems to be largely the model assumed by research integrity actors and science reformers alike, who aim to enable replication and reproducibility through procedures of explication and sharing (e.g., preregistration, data sharing, extensive method sections). The enculturational model assumes that the practice and appraisal of replication or reproducibility is part of an (epistemic) culture of the research community in question (Brenninkmeijer et al., 2025; Collins, 1975; Knorr-Cetina, 1999; Travis, 1981). In reaction to Collins, Chen states that the prototype of an (ideal) experiment replication is to a certain degree shared within a community as a form of life (Chen, 1994; also see Ulpts & Schneider, 2024). Accordingly, crucial components of what constitutes a replication and how to enact as well as appraise it remain tacit (Collins, 1975, 1985; Brenninkmeijer et al., 2025).

At least since his book *Changing Order* the role of tacit knowledge in replication practices is well established (Collins, 1985). Collins (1975, 1985) work on tacit knowledge builds on the work from Polanyi (1958). Tacit knowledge is not the opposite of formal knowledge, but more like embodied skills and implicit knowledge that is embedded in and acquired through routine practice. Such tacit knowledge is more or less shared within a specific scientific community or group depending on processes of socialisation, like PhD education (Collins, 2010; Collins et al., 2023; Delamont & Atkinson, 2001). For replication this is relevant because it means there will be aspects of research

practices that are not explicated in reports that other researchers could use like instructions to replicate a study. This even means that it can be impossible for an original researcher to communicate or explain (e.g., in writing) what the necessary and sufficient conditions are or how to get a method, procedure, or measurement to work properly (Brenninkmeijer et al., 2019; Collins, 1975, 1985; Travis, 1981), researchers might also simply not be aware of certain decisions they make and practices they engage in (van Drimmelen et al., 2024). Collins (1985) describes a situation where a research team was unable to replicate the operation of a laser. They were only able to get it to work through personal interactions with and involvement of the original researchers. Thus, only once they together conducted the practice of getting the laser to work were they able to replicate it (also see Shapin & Schaffer, 1985). Hence, whether the original researcher and the replicator are members of the same scientific community and to what degree such tacit knowledge is and can be shared within a community (between the originator and a replicator) can critically influence whether something can be replicated or not and by whom (Collins, 1985; Shapin & Schaffer, 1985).[56] Followingly, in his work on replication in memory research Travis (1981) states that sharing and publishing more information did not noticeably help with getting replications to work, because of the necessity of tacit knowledge for experimental performance (also see Shapin and Schaffer, 1985).[57] Moreover, putting current focus on replication rates into context (see e.g., OSC, 2015; Tyner et al., 2026), acceptance of certain knowledge claims seems to 'come about not through decisive experiments alone, or through the simple addition of positive results' (Travis, 1981, p. 26; also see Norton, 2015). The appraisal of replications seems inherently bound to notions of acceptability determined by the specific epistemic culture and background knowledge of the research community (Knorr-Cetina, 1999; Norton, 2015; Travis, 1981; also see Nosek & Errington, 2020). Although, how far explication of tacit knowledge is possible is still in dispute (see e.g., Feest, 2016). Furthermore, some researchers might not share all the information required to enable others to replicate their research to keep a competitive

[56] This is in accordance with Shapin and Schaffer's (1985) socio-historical account of replications of Boyles' air pump, with no one being able to replicate it without visually and physically witnessing trials of working air pumps.

[57] 'the accomplishment of replication was dependent on contingent acts of judgement. One cannot write down a formula saying when replication was or was not achieved. The construction of any device which could be taken as a successful copy of an existing pump was entirely dependent on direct witnessing. No one built a pump from written instructions alone, the transmission of pump-building and pump-operating skills required the transfer of people.' (Shapin & Schaffer, 1985, p. 281)

advantage (Collins, 1974; Horbach et al., 2026; also see Bourdieu, 1975 for scientists struggling with each other over the monopoly of scientific authority).

According to Collins (1991), replication is about establishing new or contested findings based on an agreement over what counts as a properly conducted series of experiments (Collins, 1991). Furthermore, he mentions a rule of replicability: 'Replicate your observations or have them replicated!' (Collins, 1991, p. 130). However, he acknowledges that this rule itself does not include rules for its own application. Which brings him to the formulation of the experimenters' regress (Collins, 1985), which goes something like this: When experiments are used to decide which result about a specific phenomenon one should belief or count as accurate one needs an experiment that is able to accurately detect that result or phenomenon. However, at the same time, to know whether the experiment is an accurate detector of such a phenomenon depends on the existence and nature of that phenomenon. In other words, the accuracy of experiments is assessed based on the outcome while the outcome is assessed based on the experiments (also see Brenninkmeijer et al., 2025). Therefore, we are trapped in a circle and have an endless logical regress according to Collins (1985, 1991, 2016). In the absence of agreement about what counts as a successful experiment, the experimenters' regress is often resolved by social negotiation and (non-)epistemic strategies of persuasion (Collins, 1985, 1991).

However, it would be reminisced to not highlight that, although he is the most prolific and well-known representative of the application of the 'empirical program of relativism' to replication (Collins, 1981), he remains convinced that, irrespective of the epistemic context, the aspiration towards replicability of a research community functions as a demarcation criterion (Collins, 2016). This stands in stark contrast to parts of current STS work about the relevance and feasibility of replication and reproducibility in the context of epistemic diversity (see e.g., Leonelli, 2018; 2022; 2023; Patton, 2024; Penders et al., 2019; 2020; Guttinger, 2020; Ulpts & Schneider, 2025).

Mulkay and Gilbert (1986) conducted interviews with biochemists working on oxidative phosphorylation about the nature and role of experimental replication. Based on those interviews, they distinguish between 'mere replication' (p.21) and 'replication through experimental variation' (p. 27). Mere replication is the close copying or experimental repetition, sometimes referred to as exact or direct replication, which was deemed less interesting by interviewees, than the exploration of own problems. According to Mulkay and Gilbert (1986), scientists find mere replication redundant, because there is a widely shared assumption among scientists that exact replication is always possible in principle. Therefore, it does not have to be enacted in practice.

However, mere replications are still seen as practices to identify doubtful findings, but not as truth finders. Furthermore, failure to transfer a finding to different labs is presented as a reason to doubt that finding and can even be sufficient for disregarding such a finding. In contrast, if a finding remains stable across experimental variations under diverse conditions it is seen as valid and genuine. Therefore, according to Mulkay and Gilbert (1986), mere replication seems not to be (an important) part of the social validation process of science. As they formulate it about the higher importance of replication through experimental variation over mere replication in actual research practice, meaning comes from 'the systematic pursuit and interpretation of experimental differences', not sameness (Mulkay & Gilbert, 1986, p. 35). Recently Guttinger (2018, 2019) and Peterson and Panofsky (2021) have come to similar conclusions with their concepts of micro-replication in the life sciences and the distinction between integrative and diagnostic replication respectively (also see Brenninkmeijer et al., 2025; Ulpts & Schneider, 2024).

Replication and reproducibility talk in the lab

During my field observations, the topic of replication or reproducibility arose occasionally in conversations between the researchers. For example, in the German biopsychology lab while observing a PhD student conducting their experiments the student and another researcher involved in the study told me that what they are doing is kind of like a replication of an older study, but not really. Furthermore, during one of the project proposal meetings in one of the neuroimaging labs in the Netherlands, I was able to observe a presentation of a planned (conceptual) replication study. Moreover, while sitting in during one of the lab meetings at the cognitive science laboratory in the UK one of the researchers mentioned using the Poetry tool in Python and how it increased the reproducibility of their research. Therefore, in the context of the conceptual confusion surrounding the replication and reproducibility terminology in the literature (see e.g., Albertoni et al., 2023; Barba, 2018; Gundersen, 2021; Matarese, 2022; Ulpts & Schneider, 2024), in the interviews I asked the researchers what the terms replication and reproducibility mean to them, as well as what the practices behind those terms do for their research. Hence, by having the researchers clarify what those terms mean to them, I thought it would be clearer to attempt to capture which practices the researchers link to which epistemic function of replication and reproducibility (Ulpts & Schneider, 2024).

Meaning(s) of replication and reproducibility in the lab

One of the most commonly provided understandings of reproducibility goes something like this: the ability to obtain the same results with the same analysis based on the data provided by the researchers of the original study. A caveat to the understanding of reproducibility that is sometimes added, is that someone else can do it, which in the literature is often referred to as the ability to get something independent of the original researcher (see e.g., Ankel-Peters et al., 2023; Collins, 1976; LeBel, 2015). One might argue that through the process of independent acts of redoing researchers seem to think that they can free scientific findings as well as procedures from the idiosyncrasies of specific researchers and labs. Such a process of transforming findings from a local occurrence into a general matter of fact can also be comprehended as the process of multiplication of witnesses, since it is about people other than the original researchers being able to reproduce or replicate findings (Shapin & Schaffer, 1985). Therefore, making something more stable, or communicating stability of one's findings to the research community, can be treated as an indicator of credibility (Shapin, 1995a).

> The reproducibility I just see more as, can I just even if I have your data, for example, or your code, or your description of what you did in the study, [...] reproduce the results that are reported in the paper? (PhD Cognitive Science, The UK).

While the last quote talks about someone else's work being reproducible, so that the researchers are able to use the data and reproduce a result, the next quote seems to indicate that reproducibility is also a standard they set for themselves and a characteristic of the research which makes it credible:

> You should be sure enough of your results to be confident that if someone tries your same experiment with the same parameters and the same specs, a similar result should come out of it. So, this is, of course, what I mean with reproducibility (PhD Neuroimaging Lab, Germany).

Some explicitly state that this notion of reproducibility is seen as something self-explanatory and a necessary characteristic of research to be credible and count. In other words, reproducibility must be given and 'has to happen' (PI, Neuroimaging Lab 1, The Netherlands). Most see it as basic to 'good research' or research credibility (Collins, 1985; Haig, 2022). The professor at the German personality psychology lab also adds that irreproducibility is treated like an indicator for 'bad research':

> Reproducibility is simply an absolute matter of course, if that is not given, then that is of course shoddy[58] (Professor Personality Psychology, Germany).

While for those researchers, reproducibility means being able to get the same or similar enough results with the same data, replicability and replication mean the ability to obtain the same or similar results with new data. In the literature this distinction between reproducibility and replicability is (in computational sciences) also referred to as the Claerbout/Donoho/Peng convention (Claerbout & Karrenbach, 1992; Donoho et al., 2009; Matarese, 2022; Peng et al., 2006).

> Replication, I guess for me is being able to produce the same pattern of results in a different set or in a different study. So, with different individuals given the same design you observe the same pattern of results (Postdoc Neuroimaging Lab 1, The Netherlands).

The researchers' conceptualisations of replication and replicability seem to be tied to a certain degree of sameness concerning the specific characteristics of the study and results in comparison to the original study. As one interlocutor put it: 'you could have different expectations about how similar they need to look' (PI Cognitive Science Lab, UK). This further complicates the appraisal, due to its inherent uncertainty, by highlighting that how similar something has to be to count as a replication could be a matter of dispute (Nosek & Errington, 2020; Sarisoy, 2025; Schickore & Hangel, 2019; Ulpts & Schneider, 2024). This hints towards something like the experimenters' regress (Brenninkmeijer, 2025; Collins, 1985, 2016), where lack of agreement might require non-epistemic strategies for resolution. However, at the same time, in the context of this distinction, replication and replicability is described as something more complex and complicated to achieve than mere (computational) reproducibility.

> With replication of course it's trickier. So, ideally you would also like to have results replicated, but there, of course, due to the stochastic nature of the universe there can be situations where things don't replicate (PI Neuroimaging Lab 1, The Netherlands).

> Something can be reproducible but not replicable. So that can happen that you have a study that did everything right and found an effect but then if you do it the same that somehow you can't find it (Postdoc Neuroimaging Lab 2, The Netherlands).

[58] Original in German: ‚Reproduzierbarkeit ist einfach eine absolute Selbstverständlichkeit. Wenn das nicht gegeben ist, ist es, also, aber das ist natürlich abzubilligst.'

This indicates that a failed replication is not necessarily seen as a signal for something being wrong or that there was some kind of wrongdoing by the original researchers. Responsibility for failed replication is here not allocated to researchers, but to the nature of research and the subject of investigation (see e.g., Guttinger, 2020; Penders et al., 2019, 2020; Ulpts & Schneider, 2025). The point about the 'stochastic nature of the universe' touches on features of the nature of the subject of investigation that have been highlighted as affecting replication and reproducibility, like for instance its plasticity or interactivity with the environment and the investigators as well as the dependence on time, sometimes called historicity (Guttinger, 2020; Penders et al., 2019,2020; Ulpts & Schneider, 2025; see Gergen, 1973 arguing that social psychology can be treated as a history rather than a science for such reasons). In other words, similar to previous research these remarks suggest an inherent uncertainty regarding the practice and assessment of replication (see e.g., Collins, 1985, 2016; Hangel & Schickore, 2017; Peterson & Panofsky, 2021; Schickore & Hangel, 2019; Whitley, 1984). Concerning terminology, some also go further in detail and add qualifiers like computational, conceptual, exact, direct, real, and pure. Hence, researchers outline different types of replication (see e.g., Derksen & Morawski, 2022; Gundersen, 2021; Matarese, 2022; Ulpts & Schneider, 2024).

> I call that direct replication, exactly copied/imitated replication or as exactly as possible copying/imitating replication [...] That I'd always test. That has to be fulfilled in any case. That is nearly like the reproducibility. That means the same done again in a different country in a population as it was specified in the original paper[59] (Professor Personality Psychology, Germany).

This terminology is common in social and psychological sciences (see e.g., Derksen & Morawski, 2022; Hudson, 2023; Schmidt, 2009) and direct replication, which is the redoing of a study the same or as similar as possible to the original based on the description provided in the original publication, is here likened to reproducibility. Thus, to still count as a direct replication, deviations from the original study in the replication study seem to be justified by the description of the original study. Such references to the role of the

[59] Original in German: ‚für mich sozusagen sowas, ich nenne das direkte Replikation, genau nachgemachte Replikation oder möglichst genau nachgemachte Replikation, das wäre das erste, was ich immer testen würde. Das muss auf jeden Fall erfüllt sein. Das ist fast sowas ähnliches, wie diese Reproduzierbarkeit. Das heißt, ich, ähm, dasselbe nochmal gemacht in einem Land, in einer Population, wie sie spezifiziert wurde, in dem ursprünglichen Paper. Wenn da steht, die Ergebnisse gelten nur von Deutschland, dann natürlich eine deutsche Stichprobe, in einem deutschen Lab, mit genau den gleichen Materialien, mit genau den gleichen Sachen, mit einer neuen Stichprobe, kommt wieder raus. Das ist die direkte Replikation.'

reporting of studies as instructions for (direct) replications further links to Collin's (1975) notion of the algorithmic model of replication (also see Brenninkmeijer et al., 2025). Similar to reproducibility this scholar also judged direct replication as being a necessary condition for credible research and findings. The professor of the German personality psychology lab was the only one I talked to, who goes so far as stating that when their work is not directly replicable, that would justify a commentary or even a retraction.

> That is the direct replication. […] That is the expectation (standard) I set for myself. I want to make sure that if someone redoes (imitates) the exact same thing that it also comes out. Otherwise, I think it would then be important to actually write a retraction, at least a commentary, but actually also a retraction to a paper if it, for whatever reason, is not given[60] (Professor Personality Psychology, Germany).

Although, despite the apparent agreement regarding how researchers understand and use these terms, some also use them exactly the other way around, with replication or replicability referring to the ability of obtaining the same or similar enough results with the same data and reproducibility as something broader (see e.g., Barba, 2018).

> A reproducible effect would be […] you could maybe change things about the study but keeping the kind of broad spirit of it. […] you generalise it. So, we replicated that study and now we want to reproduce it in this other population (PhD Cognitive Science, The UK).

> Reproducibility can also be even when you have small deviations you still see the same pattern of results or not so it can also be a bit more general versus

[60] Original in German: ‚Und da gibt's, ja, das ist jetzt als sprachgebrauch vielleicht nicht ganz so klar, aber für mich sozusagen sowas, wie ich nenne das direkte Replikation, genau nachgemachte Replikation oder möglichst genau nachgemachte Replikation, das wäre das erste, was ich immer testen würde. Das muss auf jeden Fall erfüllt sein. Das ist fast sowas ähnliches, wie diese Reproduzierbarkeit. Das heißt, ich, ähm, dasselbe nochmal gemacht in einem Land, in einer Population, wie sie spezifiziert wurde, in dem ursprünglichen Paper. Wenn da steht, die Ergebnisse gelten nur von Deutschland, dann natürlich eine deutsche Stichprobe, in einem deutschen Lab, mit genau den gleichen Materialien, mit genau den gleichen Sachen, mit einer neuen Stichprobe, kommt wieder raus. Das ist die direkte Replikation. Ähm, das muss auch erfüllt sein, das sollte man prüfen und jede, Forschung, die ich veröffentliche, muss ich zumindest mal überzeugt sein, dass sie direktm replizierbar ist. Das ist die Erwartung, die ich an mich stelle. Ich möchte alles dafür sicherstellen, dass wenn jemand genau das Gleiche nachmacht, dass es dann auch rauskommt. Ähm, ansonsten glaube ich auch, das wäre was, was man dann tatsächlich, das wäre wichtig, eine Retraction zu schreiben, ein, äh, mindestens ein Comment, aber eigentlich auch eine Retraction zu einem, zu einem Paper, wenn das dann, warum auch immer, wegen irgendwas nicht so ist. Aber mein Anspruch wäre, alle meine Forschung lässt sich direkt replizieren.'

> replication is much more tied to the design and the specific choices that have been made (Postdoc Neuroimaging Lab 1, The Netherlands).

Confusing terms

Some researchers even hint towards disagreements and diversity concerning the use of terminology and meanings of reproducibility and replication, something like a conceptual confusion or proliferation of terminology across and within disciplines (see e.g., Albertoni et al., 2023; Barba, 2018; Gundersen, 2021; Matarese, 2022; Plesser, 2018; Ulpts & Schneider, 2024). This includes the circumstance that 'they are used differently in different corners'[61] (Postdoc Personality Psychology, Germany) and that 'some people use them interchangeably' (PhD Cognitive Science, The UK). Then, when trying to distinguish the terms, there are also utterances indicating some form of confusion about which term is used for what type of practice ending up to 'always get those two mixed' (Postdoc Neuroimaging Lab 1, The Netherland). However, at the same time others also make it very clear that there is no problem at all to clearly define these terms within the confines of their discipline and in the context of their research:

> In psychological, social science research it is completely unproblematic to define[62] (Professor Personality Psychology, Germany).

Which could mean that there is some kind of tacit agreement between researchers that are part of the same research communities, due to a shared socialisation process, eliminating the need for explicating their meaning (Chen, 1994; Collins, 1975). Or researchers just use terms like replication and reproducibility as indicators for credibility or trustworthiness of the research and findings without looking behind the practices underlying these terms. Additionally, a conviction of agreement about replication and reproducibility terminology can also suggest a lack of awareness about the proliferation of terminology. However, when distinguishing between types of replications like conceptual and direct replication, the professor declares a lack of clarity in language:

> And there the language usage is maybe not that clear (Professor Personality Psychology, Germany).

[61] Original in German: ‚Ja, die werden ja so in verschiedenen Ecken unterschiedlich gehandhabt.'

[62] Original in German: ‚in psychologischer, sozialwissenschaftliche Forschung vollkommen unproblematisch zu definieren.'

Furthermore, when terms like reproducibility and replication are used in the calculation of replication rates like in large scale replication projects (see e.g., OSC, 2015; Cova et al., 2021; Errington et al., 2021; Tyner et al., 2026), different practices can turn out to be hidden behind one single numeric indicator, obfuscating the existing diversity in practices behind the same terminology resulting not just in a diversity of replication rate metrics, but also in potentially invalid replicability rates (see e.g., Bak-Coleman & Devezer, 2024).

Function of reproducibility and replication

The topic of a replication crisis comes up when researchers are asked about the value and meaning of replication for their own research (Brenninkmeijer et al., 2025). Hence, one thing interviewees mention regarding why replication and reproducibility are or have become such important topics in recent years is because of the perceived replication and reproducibility crisis in psychology, neuroscience and related research areas (Brenninkmeijer et al., 2025; Peterson & Panofsky, 2023; Romero, 2019).

> Because in our field a lot of results cannot be replicated[63] (PI Neuroimaging Lab, Germany).

> Yeah. Of course, these are big topics also in the whole community because we do have a replication crisis (PhD Neuroimaging Lab 2, Netherlands).

However, there is also criticism about such a diagnosis of a replication crisis state in psychology (see Fanelli, 2018; Feest, 2019):

> Replication, is actually something smaller than the current discussion about replication, to right away declare a crisis or disease. So, crisis and disease maybe comes from the naïve belief that 100% of all you read can immediately without any effort be shown in all disciplines. And that's just not the case, everyone involved in research actually knows that, but Psychology is very enthusiastic about this, especially about aspects where it doesn't work, so well. […] This discussion isn't unimportant, but I think we have to be careful that we don't dedicate too many research programs to this discussion, because then it distracts again from important substantive discussions[64] (Professor Biopsychology, Germany).

[63] Original in German: ‚Weil in unserem Feld total viele Ergebnisse nicht replizierbar sind.'
[64] Original in German: ‚Und das glaube ich ist ein wichtiger Aspekt der Replikation. Die eigentlich, sozusagen kleinere Brötchen als die aktuelle Diskussion über die Replikation das gleich wieder als Krise und Krankheit zu erklären also Krise und Krankheit ist vielleicht ist naiver Glauben sozusagen in jeder Wissenschaft kann 100 Prozent von dem, was man liest sofort und ohne weiteres wieder ohne Aufwand gezeigt werden. Und das ist einfach nicht so, das Wissen eigentlich alle die Wissenschaft machen und aber die Psychologie begeistert sich sehr stark dafür, gerade wenn es dann um sozusagen Aspekte geht wo das nicht so gut

Indicating that while important, replication is a complex issue that can be treated in an overly simplistic manner, enacted by a hyperfocus on replicability and reproducibility issues in the context of a moral panic associated with the replication crisis narrative (Baker, 2016; Bartscherer, 2026; Bartscherer & Reinhart, 2025; Earp & Trafimow, 2015; Haig, 2022; Penders, 2024; Penders et al., 2020; Peterson & Panofsky, 2023; Rubin, 2023). This seems to be seen as, at least somewhat, troublesome because replication is one among many issues or aspects of research in psychology that have to be considered for establishing and assessing credibility in the context of limited resources and time to conduct research (see Guttinger, 2020; Latour & Woolgar, 1979; Penders et al., 2019; Peterson & Panofsky, 2021; Schickore & Hangel, 2019). Hence, this professor suggests that it is essential to consider what resources to dedicate and maintain a balance with other practices and criteria conducive to research integrity or the integrity of the discipline. Such comments also complicate the current reliance on replication rates for a crisis diagnosis (see e.g., OSC, 2015), because comparability of replication rates will also be influenced by what can actually be excepted based on the characteristics of the research and research subject in question (Guttinger, 2020; Leonelli, 2018; Penders et al., 2019; Ulpts & Schneider, 2025). While the perception or narrative of a replication crisis seems to be part of a cultural context in which issues of replication become more salient for members of the research community, the researchers also provide explicit epistemic functions and motivations for practicing replication and reproducibility.

Providing a stable foundation for incorporation

A reoccurring motive behind reproducibility and replication is to assess how solid or stable a finding is. This function of replication as a way of testing and establishing how substantive a finding or study is seems again linked to the notion of independence of the replicator from the originator as a crucial ingredient for trust in previous research, that someone else somewhere else can do it too (Ankel-Peters et al., 2023; Collins, 1976; LeBel, 2015).

> An experiment that can only be replicated at the lab where the effect was originally found is probably bullshit (Postdoc Neuroimaging Lab 1, The Netherlands).

The underlying motivation behind the establishment of stability through replication presented is that researchers can build on previous research to potentially incorporate it into and proceed with their own research (Mulkay &

---

gelingen der Psychologie [...] Also das ist also insofern ist die Diskussion nichts nicht unwichtig, aber ich glaube man muss aufpassen dass wir nicht zu viel Forschungsprogramme auf dieser Diskussion aufbauen, weil sie dann wieder abzieht von wichtigeren inhaltlichen Diskussionen.‘

Gilbert, 1986; Peterson & Panofsky, 2021; Whitley, 1984). They seem to see reproducibility and replication as practices that tell them whether previous research is solid enough to build up on it and as indications whether there really is something to build on. The next quote, for instance, indicates that reproducibility functions like a mechanism that communicates to the researcher that something like the code runs well, works, and can be used and build on as a component in their own research. Indicating, a motive of reproducibility similar to what Peterson and Panofsky (2021) term integrative replication, where specific components of previous research that might be useful for their own research are (adapted and) incorporated for one's own purposes (also see Guttinger's, 2019 notion of micro-replication). Hence, instead of policing, repairing or diagnosing error and bad research as part of a moral economy of repair (Penders et al., 2020), the responses indicate a rather integrative motive (Guttinger, 2019; Peterson & Panofsky, 2021).

> If a code is reproducible, then it means I can take it and use it for how I want to (PhD Cognitive Science, The UK).

This integrative motive is also visible in responses suggesting replication is a practice that also tells the researcher that they can trust their own research, if they see similar outcomes as previous research. Therefore, it is not just seen as a backward-looking procedure to check or correct previous research in the scientific literature (Flis, 2022; Penders et al., 2020; Peterson & Panofsky, 2021), but also as a forward-looking mechanism that indicates to a researcher whether they can trust their own practice, build-up, and share their findings with the world (Peterson & Panofsky, 2021).

> Results that fit very well with previous studies, that report exactly similar effects, similar things. And that, for example, for me personally was pretty cool, to see that back then, and has also a bit strengthened my trust in the results somehow[65] (Postdoc Neuroimaging Lab, Germany).

Accordingly, validity of findings seems to be conceptualised as a form of independence of the local conditions of knowledge production and a degree of stability that withstands scrutiny.

> In a certain sense, once you have achieved something, a result, and then communicated it, you have the expectation that it will not vanish into thin air already

[65] Original in German: ‚Ergebnisse generiert, die sehr, sehr gut zu Studien passen, die genau ähnliche Effekte, ähnliche Sachen berichten, und das hat zum Beispiel, also für mich persönlich war das ziemlich cool, damals das zu sehen, und hat auch so ein bisschen das Vertrauen in die Ergebnisse gestärkt irgendwie, und in auch die Sachen gestärkt.'

during their communication, but that it will have a certain validity[66] (Professor Biopsychology, Germany).

Part of this process of ensuring that something is actually there and works seems to be the aim to ensure usability and also to reduce the potential for waste, for instance, in the form of uncoordinated repetition of effort. We could say that this is a process of preventing a form of replication that is seen as redundant and a waste of time and resources (NAS, 1992). A PhD student in the UK at a cognitive science lab mentioned that reproducibility can ensure that something can be (re)used and so in that sense efficiency could be increased while waste could be reduced because 'you can rely on it without wasting time [...] reinventing the wheel' (PhD Cognitive Science UK). It might be seen as a way to contribute to a system where researchers do not work past each other.

A PhD student at the personality psychology lab in Germany, who is involved in metascience, also mentions this efficiency and waste reduction function of reproducibility as a framing that makes the issue of reproducibility understandable and shows its relevance even to the wider general public. Hence, efficiency and waste reduction seem to not just be epistemic functions of replication and reproducibility in their research, but also useful as messages that aim to persuade (lay)people or stakeholders of the importance of reproducibility.

> Reproducibility is nice from a scientific perspective, but waste is also a thing that is immediately comprehensible for lay people too. [...] waste of the scientists' time, waste of the participants time and of course waste of money [...] It's so much waste of money and that is tax money that is the people's money (PhD Personality Psychology, Germany).

From such a metascientific perspective this also seems to suggest a need to not just convince the scientific community of the importance of reproducibility, but according to this metascientist even the general public or other non-scientific actors need to believe that it is a crucial component for an efficient and well working science. Further explaining waste as financial burden due to redundancies resulting in the waste of taxpayers' money. Hence, here, to borrow from Shapin (1995a), the concern about reproducibility is located as part of

[66] Original in German: ‚Genau, also es ist auf der einen Seite natürlich klar, wenn wir davon ausgingen, dass nichts von dem, was wir machen, replizierbar ist, dann bräuchte man den Mund nicht weiter Aufmachung. Wenn das alles nur Zufalls- und Eintagsbefunde wären. Man hat in gewisser Weise schon, wenn man was, ein Ergebnis rausgekriegt hat und das dann auch kommuniziert, den Anspruch, dass es nicht während der Kommunikation schon sich in Luft auflöst, sondern dass es eine gewisse Gültigkeit hat.'

the active communication in two vectors of credibility, the scientific community and the general public.

Stories of their own practice

Besides a general sense of difficulty and uncertainty associated with replication and reproducibility (Collins, 1975; 1985; Schickore & Hangel, 2019), when talking about their own practice during the interviews, it turns out that most researchers are hesitant towards the practice of replication in the form of copy-pasting or imitation-like redoing of previous research. This was one of the reasons why in Mulkay and Gilbert's (1986) study the biochemists called such practices 'mere replication' (p. 21). In the research integrity literature and the metascience literature this type of replication practice is often referred to as a direct or exact replication, however, during our interviews at the same time interlocutors tend to call such a type of replication 'real', 'pure', 'full on' or 'actual' replication. They seem to think that everything that is not a one-to-one copy or as close as possible, concerning all the relevant aspects, to previous research, is not actually a 'real' replication.

> A real replication is, of course, if we really would do the same again[67] (Postdoc Neuroimaging Lab, Germany).

Although pure or real replication is described as not often practiced or even seen as less valuable for their own research, their use of terms like pure, real or actual does suggest an understanding of replication that could explain why what they themselves then do as 'partial' replications are not actually real replications for them or declared as such (see below). This might also be one of the reasons why in studies where metascientists attempt to capture the amount of (direct) replication studies in the literature the numbers are so low and might underestimate the actual amount of replications (see e.g., Clarke et al., 2024; Makel et al., 2012), because if it is not a real replication the researchers themselves probably do not report it as such in their publications (see e.g., Guttinger, 2018; Haig, 2022; Leonelli, 2018; Ulpts & Schneider, 2024). Another reason might be as Lubega et al. (2023) suggest, that certain replication studies are kept internal to the specific labs and not published at all, similar to many pilots (Hangel & Schickore, 2017). As outlined in the preregistration chapter, according to the scholars there are cases where the published study is actually a replication of a lab internal pilot, but the published version does not communicate that to the community. As one PhD in a lab in Germany put it when reserving the term replication for a particular type of study that redoes

[67] Original in German: ‚naja also eine echte Replikation ist natürlich, wenn wir wirklich dasselbe nochmal machen würden.'

research as much as possible and as close as possible compared to an original study:

> Everything else in my opinion, there you have to say: here we have oriented on it/ based our study on it, but we also changed things. And then I would say, it just isn't a real replication anymore, then it is rather an extending/advancing study[68] (PhD Biopsychology, Germany).

Since most scholars said that they rarely conduct such replications, I also inquired into when this form of replication would be something they would consider for their own research. To which they responded that when reported findings are counterintuitive, surprising or contradict the current state of the literature some see replication as valuable and something they would do to see whether it is credible (Mulkay & Gilbert, 1986).

> When someone finds something very surprising [...] you would want to see it twice before you trust it (Postdoc Neuroimaging Lab 1, The Netherlands).

This aspect of wanting something replicated when it is surprising in the context of the existing scientific literature, is what happened in the Obokata 2014 misconduct case that was uncovered because researchers in the community were startled by their claims and tried to reproduce it and could not find the results from the previous study (Meskus et al., 2018; Resnik & Shamoo, 2017).[69]

However, for interpreting the researchers' responses it is important to consider the difference between what they do and what they say they would do or like to do, there is always the confounding factor of social desirability and the dissonance between what scientists say they do or would do and what is actually enacted in scientific practice (see e.g., Anderson et al., 2007; Mulkay, 1976). Moreover, when they talk about their own research practice the interviewees presented a form of replication that one might call 'partial' replication. This form of replication is similar to the type of replication practice that Mulkay and Gilbert (1986) identified and called 'replication through experimental variation' (p. 27), and integrative replication as described by Peterson and Panofsky (2021) as well as Guttinger's (2018, 2019) notion of micro-replication in the life sciences. In danger of worsening the conceptual confusion through terminological proliferation, I here call it partial replication to

[68] Original in German: ‚Alles andere, da muss schon meiner Meinung nach dann gesagt werden, hier wir haben uns daran orientiert, aber wir haben auch Sachen verändert und dann würde ich sagen, ist es eben keine richtige Replikation mehr, dann ist es eher eine weiterführende Studie, meiner Meinung nach.'

[69] Also see the recent controversy concerning claims about room-temperature superconductivity and the role of replications in their eventual retractions (Garisto, 2023).

capture the language of the researchers, who talk about taking parts of previous research and replicating it as parts of their own research. Neither Guttinger (2018, 2019) nor Peterson and Panofsky (2021) do discuss or even cite the work from Mulkay and Gilbert (1986). Therefore, one could say that they are unknowingly reinventing the wheel, by actively forgetting preexisting relevant work on the topic of replication that might not just have aided their empirical investigation, but also their critical analyses of the practices and functions of replication across the sciences. However, Brenninkmeijer and colleagues (2025) make the connection.

> What we therefore usually try to do is build up research. First there is a component of it that is just a replication of something else and then we add on that. […] then you start with that part and that also then gives more confidence to the previous findings. So, part of your research is valuable already, because it replicates previous research, but it also then, adds value to whatever else you find if it is a positive finding, or a negative finding (PI Neuroimaging Lab 1, The Netherlands).

Similarly, a researcher in a neuro-imaging lab in Germany, when talking about extending on their own work in the lab and using a protocol previously used for fMRI studies that they now employ in their TMS studies said:

> But what we of course often do is to somehow at least go one step further and try to answer a little bit different questions, but not change too much, so that we know to some extent what is happening and can learn from our previous research. […] That's what we usually do, that we conduct such a partial replication based on our previous studies where some or one part are then exactly like before[70] (Postdoc Neuroimaging Lab, Germany).

Consequently, taking parts of previous research for which it is already known what should happen and what to expect seems to be part of the previously mentioned function of replication to build a solid foundation for their own research endeavours. Simultaneously, with partial replication there is still space for implementing something new and adapting procedures to one's own research interests and questions to gain novel insights (Guttinger, 2019; Mulkay & Gilbert, 1986; Peterson & Panofsky, 2021; Schickore & Hangel, 2019).

[70] Original in German: ‚Aber was wir natürlich oft machen, ist irgendwie zumindest dann ein Step weiter zu gehen schon zu versuchen eine bisschen andere Frage zu beantworten aber nicht zu viel zu verändern das wir zumindest halbwegs noch wissen was vielleicht abgeht und aus unseren vorherigen Sachen lernen können. Also zum Beispiel eben hab Ich jetzt fast genau dasselbe design was ich im fMRT benutzt habe, habe Ich dann im TMS benutzt. […] Also so machen wir es ja meistens, dass wir so eine Teilreplikation machen aus unsere vorherigen Studie, ein bisschen was oder ein Aspekt ist dann genau wie vorher …'

However, even for those who have aspirations about computational reproducibility or direct replicability and see it as a necessity and set such high bars for themselves sometimes suggest a clash between their values or aspirations and the reality of conditions of their research. A postdoc in a personality psychology lab in Germany somewhat complexifies and complicates the issue of ensuring and checking reproducibility, when talking about the role and necessity of (computational) reproducibility, by shortly touching on how the feasibility of such aspirations and ideals can be limited (see e.g., Leonelli, 2022, 2023; Ulpts & Schneider, 2025). Afterwards the postdoc quickly explains how to compensate for or deal with the confrontations between aspiration and reality that arise from conducting research with limited time and resources.

> Computational reproducibility is that you have the data and get the same out when you run the analysis. […] If that is not given, we do not have to discuss it further, because we do not know whether it really is in the data. We ensure this by having control loops, that someone else runs the analysis and sees whether the same comes out. This takes time, but it is our duty as researchers, especially when it is doable, and the analysis does not run for 10 days on a cluster. However, even then you would build in some checks (Postdoc Personality Psychology, Germany).

The actual feasibility and the consideration of time come into play again, indicating something like a strategy (or justification) of investment when considering replication or using Open Science practices to be reproducible (Latour & Woolgar, 1979; Peterson & Panofsky, 2021; Shapin, 1995a), because even when implementing a workflow that enables replication, it can still be seen as unlikely that someone will actually use and replicate it.

> It's not because you publish all of your preliminary results or whatever that someone is going to replicate your study. There will be still a lot of studies that nobody cares about and they will not get replicated, they will not get scrutiny, they will just hang out there in the vast black that is the internet (PhD Neuroimaging Lab, Germany)

Failed replications

The aspect of feasibility or what can be expected concerning replicability of research or findings also comes up again in the context of experiencing failed replications. There is the, previously mentioned, sentiment that failing to replicate something is just part of the type of research that they are involved in and a normal occurrence, which might also suggest that the previously self-set standard of being directly replicable, as stated by the personality psychology professor, is actually quite a high standard and maybe even unrealistic in some cases or some epistemic contexts (see e.g., Guttinger, 2020; Leonelli, 2018;

Penders et al., 2019, 2020). Furthermore, it puts into doubt Resnik and Shamoo's (2017) direct link between lack of reproducibility and uncovering research misconduct cases. As mentioned above, there were other circumstances at play, like claims about triggering pluripotent stem cells contradicting the state-of-the-art scientific literature at the time and the involvement of a somewhat vigilant online community (Meskus et al., 2018; Pulverer, 2014), highlighting the context dependence of replication appraisal (see e.g., Norton, 2015; Travis, 1981).

> Reproducibility, especially in behavioural sciences, I think has become very, very difficult. And especially when you publish some empirical results, and all the sudden someone recreates the experiment under, as far as they know, about the most similar conditions or the exact conditions, and then the effect doesn't come out (PhD Cognitive Science, The UK).

Such statements about the uncertainty linked to replication can be ascribed to the dynamic, complex, and context sensitive nature of their subject of investigation. However, it might also be, because there is a lack of knowledge about what keeps an effect alive, or the knowledge remains tacit and, therefore, researchers are unaware of what needs to be adapted or taken care of to ensure that the effect remains alive under different or even assumedly same conditions in a different lab (Collins, 1975, 1985).

Another researcher also talked briefly about the role of failed replications and how to react to such cases when they themselves failed to replicate their own research. The scholar talks about a certain compromise between wanting to improve on previous research and the resulting risk of failing to replicate, because it is not completely clear which factors might affect or even 'kill' (Postdoc Neuroimaging Lab, Germany) the target effect. This, of course, complicates and adds uncertainty to their practice of partial replication. Although, at the same time failed replications are here described as a source of insight and an opportunity to learn and explore what might have happened to an effect (see e.g., Derksen & Morawski, 2024; Meirmans et al., 2025; Penders & Janssen, 2018):

> I think first you search for yourself for explanations. Why can it be? And you probably want to discuss that in a paper, exactly like other readers would then also ask themselves why is this different now? […] And I of course learn through this for my next study[71] (Postdoc Neuroimaging Lab, Germany).

[71] Original in German: ‚Ja absolut, also Ich glaub erstmal für sich selbst sucht man nach Erklärungen natürlich, warum. Warum es so sein könnte, und das willst du dann wahrscheinlich auch nachher in deinem Paper diskutieren genauso wie andere Leser sich dann ja auch fragen würden die deine Forschung vorher schon kennen, wieso ist das jetzt anders?

However, at the same time, the sense of suspicion within the research community, which I call a culture of suspicion in Chapter 7, here further seems to find its realisation, for example, as a suspected fear of being marked as irreplicable. One researcher, for instance, talked about a suspicion that some researchers' work might just not replicate and they intentionally do not enable others to replicate their work to protect themselves, because they could have a certain fear of being replicated, due to them knowing or thinking that their results may not hold up. Hence, here a scholar, in contrast to Collins' (1974) case where it was about competitive advantages between competing researchers (also see Horbach et al., 2026), voices suspicion about researchers preventing replication of their work as a self-protection mechanism against suspicion. Whether that is due to competition or other reasons remains unclear. Also again indicating that, despite previous remarks, people in the community see failed replications or irreplicability as a reason for suspicion toward the conduct and researcher behind the original study:

> I think a lot of people are simply afraid of that, because they know or think; my result can maybe not be replicated[72] (Postdoc Neuroimaging Lab, Germany).

Hence, we observe a certain ambivalence among scholars. While there seems to be some kind of stigma associated with non-replicability among some scholars, for others it is just a normal consequence of the nature of their research. In that vein, such utterances seem to indicate that failed replications or the inability to conduct a replication study can provoke suspicion or accusations of questionable or bad conduct towards the original researchers. One potential reason for failed replication that are again linked to a sense of suspicion and shifting responsibility towards the original researchers themselves is the recurring theme about the role of replication and replicability in relation to not necessarily malice, but error and bias (Frank et al., 2026). Which is part of characterising the researcher as fallible (Flis, 2019; Morawski, 2020). However, it also indicates that instead of seeing failed replication as a sign of misconduct, it can just be honest error (see e.g., Frank et al., 2026):

> I think that you can make many mistakes, and everybody can (PhD Neuroimaging Lab 2, The Netherlands).

[...] Und dadurch lerne Ich natürlich wieder für meine neue Studie [...] Ja aber man weiß es nicht man versucht eine Sache zu verbessern und macht eine andere Sache dadurch schlechter. So ist das. Auch da ist der Kompromiss immer wieder.'

[72] Original in German: ‚Ich meine, ich glaube, viele Menschen haben einfach total Angst davor, weil sie wissen oder sich denken, okay, mein Ergebnis kann man vielleicht nicht replizieren.'

Unfortunately, like in the case with successful replications it remains unclear what exactly constitutes a failed replication, and I did not further inquire into this, and the researchers did not provide more detail concerning what constitutes a failed replication or how to deal with it. Which would be an interesting opportunity for future research, because it is also an area of ongoing dispute in the literature, what constitutes a successful and failed replication (see e.g., Nosek & Errington, 2017; Rubin, 2023; Sarisoy, 2025), and how failed replications might not just be received differently across communities, but also serve varying purposes (Meirmans et al., 2025).

Reproducibility and replication and a perceived demand for novelty

Similar to previous research into the topic of reproducibility and replication in the context of actual research conduct under real life conditions (see e.g., Horbach et al., 2026; Mulkay, 1984; Mulkay & Gilbert, 1986; Schickore & Hangel, 2019), some responses indicate that researchers have to navigate the aspiration or ideal of conducting replication studies in the context of a system or close environment where they perceive such practices as less or not wanted (also see Lilienfield, 2017; Mulkay, 1984; Nosek et al., 2012; Romero, 2018; Vachon et al., 2021). This specific presented conflict between aspirations to conduct replication studies and the research system or even their own research institute preferring novel, original, ground-breaking, and high-impact research publications, is perceived as pressure at different levels and from various bodies. Hence, for the researchers the foreseen return on investment for replication and reproducibility within the current system might sometimes just not be worth the effort or this is how they rationalise not conducting replication studies (see e.g., Elizondo & Kaltenbrunner, 2024; Horbach et al., 2026; Latour & Woolgar, 1979; Peterson & Panofsky, 2021; Schickore & Hangel, 2019).

> I find that extremely annoying at the [institute], this expectation that everything always has to be excellent and high-risk and top-notch and cutting-edge. That is simply not realistic at all. The biggest ones usually proceed in small steps, and I don't have to think for every study: 'Ah, this has to be published in Nature.' I also don't have the impression that this is the culture that is always pursued, but I think some working groups are pushing it too hard in that direction, and that contradicts the idea that you want to do a replication study if possible and make it transparent and open[73] (PI Neuroimaging Lab, Germany).

[73] Original in German: ‚Ich glaube insgesamt, also ich denke unser Institut so methodenmäßig und Infrastruktur ist es schon top, da hast du schon mehr als an den meisten anderen Institutionen. Was denke ich, sich noch mehr durchsetzen muss, ist weg von der Kultur, dass

> And when you somehow do a replication … Although, by now that is more endorsed, but actually if you do the same again, you're probably anyway not getting a research grant. Because… what's the novelty in the whole thing? Then it is rejected anyway. Partly it still doesn't make that much sense, that it wouldn't do much for your career or your research[74] (Postdoc Neuroimaging Lab, Germany).

This (publication) bias on the side of institutes, journals, and funders towards novelty is a reference to what is called perverse incentives (see e.g., Edwards & Roy, 2017; Ferguson & Heene, 2012; Lilienfield, 2017; Rosenthal, 1979; Stephan, 2012). Simultaneously, as was already indicated by talk about a replication crisis, there also seems to be the feeling that the issues surrounding replication and reproducibility are more prevalent in peoples' minds, but actual change in the system seems to be perceived as missing: 'Even if it's talked about more and more, it is still not that valued' (Postdoc Neuroimaging Lab 1, The Netherlands).

The professor of the German personality psychology lab also talked about the hardship of trying to get a failed replication published as well as how other more senior scholars always warned not to bother with replications, due to that hardship inherent to the system. Furthermore, knowledge about which findings are irreplicable is described as kept at conferences as part of an 'invisible college' (Crane, 1972). Going on to shortly elaborate on a case where they failed to replicate a study that was already published in a journal and how the prophecies of the senior researchers at the time became true. When they submitted the failed replication, they were told by the editors that it is their obligation to explain why it failed, because the other study was already published and assumed to be something like established truth:

> I was told as a doctoral student, under no circumstances should you replicate anything; if it doesn't come out, you won't get it published anywhere. […] And only later did I try to do it myself, with a doctoral student […]. And now this is

---

es immer nur High-Risk-Forschung sein muss. Also das finde ich am [Insititut] extrem nervig, dieser Anspruch, dass alles immer Exzellenz sein muss und High-Risk und Top-Notch und Cutting-Edge. Also das ist einfach überhaupt nicht realistisch. Die größten, der größte, also meistens geht es halt kleinschrittig und ich muss doch nicht bei jeder Studie mir überlegen, ah das muss in Nature kommen. Ich habe auch nicht den Eindruck, dass das die Kultur ist, die immer so gefahren wird, aber von einigen Arbeitsgruppen finde ich, wird das schon zu sehr in die Richtung gepusht und das steht ja dem gegenüber, dass man möglichst auch mal eine Replikationsstudie machen will, es transparent und offen macht.'

[74] Original in German: ‚Und wenn du irgendwie eine Replikation… Das wird zwar inzwischen mehr gefördert, aber eigentlich, wenn du nochmal dasselbe machst, kriegst du wahrscheinlich sowieso keinen Forschungsgrant. Weil du… Was ist jetzt die Novelty in dem Ganzen? Dann wird's eh abgelehnt. […] macht noch teils noch nicht so viel Sinn, dass es dir für deine Karriere oder deine Forschung dann so viel bringt.'

> exactly what happened, exactly what I was always told would happen. Yes, but it's published in [Journal name] and that makes it clear. And we have to explain why we can't get it out. It says so in the editorial. We have to deliver, because the other thing has been published and there's no chance. It was possible because the publication journals changed a bit. We then got it published in [journal name]. But when I was a doctoral student, […] it was completely impossible. […] you only said it at conferences: 'but the findings are not replicable.' […] Only those who went to conferences and talked to people even found out about it[75] (Professor Personality Psychology, Germany).

Related to this publication bias, people seem to behave as if the act of publication signals the end of the critical process of assessing and reflecting on a specific matter of concern and it is treated as an established matter of fact, sealed by the act of publication (Latour, 2004). Therefore, there seems to be a higher threshold for post-publication criticism that affects how especially failed replications are received, since they potentially challenge published findings. Such frustration with the common journal publication model in academia has led to calls and implementation of reform towards the publish-review-curate model of publication (Corker et al., 2024), for example, in the form of post publication peer review platforms like PubPeer (see e.g., Meskus et al., 2018). These problems of trying to publish (failed) replications of already published work reminds one of stories like Ritchie and colleagues' (2012) attempt at publishing a replication study that failed to find evidence for precognition as the

[75] Original in German: ‚ich glaube, lange Zeit diese Replikation auch nicht als besonders wichtig angesehen wurde, auch nicht, auch nicht publizierbar war. Mir wurde noch als Doktorand/ Doktorandin gesagt, auf keinen Fall was replizieren, wenn das nicht rauskommt, das kriegst du nirgendwo veröffentlicht. Ja, verschwende nicht deine Zeit damit, um ihn irgendwo gar nicht zu veröffentlichen oder in einem drittklassigen Journal oder fünfklassigen Journal vielleicht irgendwann, aber selbst das kriegst du nicht hin, weil du hast immer irgendwas nicht falsch gemacht. Die finden immer irgendeinen Grund. Und erst später habe ich das dann versucht, auch selbst zu machen, mit [Doktorand/ Doktorandin], der/die ist auch aus der Wissenschaft rausgegangen. Da haben wir das dann tatsächlich versucht, mit großen Stichproben den Befund zu replizieren. Und jetzt genau das passiert. Genau das, also genau das, was mir immer angekündigt wurde. Ja, aber es ist in [Journal] veröffentlicht. Und damit ist es klar. Und wir müssen erklären, warum wir es nicht rausbekommen. Steht so im Editor drin. Wir sind in der Bringschuld, weil das andere ist veröffentlicht und es ist keine Chance. Es war dann möglich, weil sich das Publikations journals ein bisschen geändert hat. Wir haben es dann damals in [Journal], veröffentlicht bekommen. Aber zu der Zeit, als ich noch Doktorand/ Doktorandin war, gab es das gar nicht. Da gab es nur die Psychologie-Journals und damals war das völlig unmöglich. Das heißt, man hat alle Leute, die irgendwas, man hat es nur auf Konferenzen gesagt. Aber die Befunde sind ja gar nicht replizierbar. Subliminale Priming, das ist natürlich nicht replizierbar. Das stand nirgendwo. Nicht mal im Blogpost. Das gab es nicht. Das gab es nur auf, nur wer auf Konferenzen gegangen ist und mit den Leuten gesprochen hat, hat das überhaupt erfahren.'

now infamous paper by Bem (2011) did, whose publication in one of the leading social psychology journals was one of the factors leading to claims about a replication crisis in (social) psychology (see e.g., Romero, 2019; Wiggins & Christopherson, 2019). The Journal of Personality and Social Psychology (JPSP), which published Bem's original study, did generally not publish replication studies at the time and, like in the story of the researchers here, they eventually also got it published at another journal (Ritchie et al., 2012).

However, in the current science system scientists themselves are largely involved in the assessment processes that decide who gets hired, published and funded. Therefore, the internal-external distinction often underlying the act of 'blaming the incentives' or claiming powerlessness in the context of 'evaluation regimes' can be problematised (Cramer & Reinhart, 2026; Robinson-Garcia et al., 2023; Schneider et al., 2021; Yarkoni, 2018). Furthermore, as the prevalent practice of partial replication already suggests, it seems not just to be an external systematic influence incentivising or demanding novelty and originality which might work against the practice of (real) replications, researchers themselves also tend to want to do and learn something novel (Knorr-Cetina, 1981, 1999). As the next quote, arguing for why to do a partial replication instead of a direct replication illustrates: "I mainly really wanted to also learn something new[76] (PhD Biopsychology Germany).

The professor of the German biopsychology group also puts the necessary investment for a direct or pure replication in relation to personal research interests and whether such an investment of the effort to redo something as close as possible to an original study is worth it for them:

> There is always the question of how much effort is it worth for me? How much am I willing to put into a replication that's too narrow without sacrificing the freedom to actually research what I want to research?[77] (Professor Biopsychology, Germany).

Interestingly, a psychology postdoc involved in metascience in Germany seems to go in the opposite direction and shortly discussed that replication alone might actually not be enough, suggesting that for it to work properly it needs to be combined with (other) Open Science reform procedures, to form

[76] Original in German: ‚was sich gut publizieren lässt und nicht, hat vielleicht schon auch im Hintergrund ein bisschen mit reingespielt, aber hauptsächlich wollte ich tatsächlich auch was Neues lernen. […] Und das war sozusagen der Hauptgrund, also wirklich, dass mich interessiert hat, was Neues herauszufinden und nicht, dass ich mir gedacht habe, ne, eine Replikation lässt sich halt nicht so gut publizieren.'

[77] Original in German: ‚aber dann ist immer die Frage, was wie viel Aufwand ist mir das wert sozusagen? Also wie viel stecke ich rein in eine zu enge Replikation, ohne mir die Freiheit zu nehmen, eigentlich das zu beforschen, was ich beforschen will.'

something like, what Derksen and Morawski (2022) have, in reference to Law (2004), called a 'method assemblage' (p. 1497). The aim of such an Open Science method assemblage is to correct and protect research against bias, errors and misbehaviour from multiple angles and layers, by addressing different phases and components of the research workflow. Derksen and Morawski (2022) identify such a method assemblage as part of the Open Science movement 'for the production of reproducible science (Munafò et al., 2017, p. 1)' (p. 1497):

> I know certain fields, where I have the feeling, replications were not enough. […] My non-replications had some impact there, but surely even higher, because they were preregistered and as registered reports, they would have probably been even better[78] (Postdoc Personality Psychology, Germany).

In the context of the experimenters' regress resulting from the uncertainty surrounding the practice and appraisal of replication, especially in the biomedical and social science (Feest, 2016, 2019; Peterson & Panofsky, 2021), it seems as if the researcher is describing a process of moving closer towards algorithmic enactments of replication (Collins, 1975), enabled buy procedural interventions like preregistration or registered reports. This is assumed to make the claims more convincing to the research community, by persuading them of their credibility (Shapin, 1984, 1995a). OS practices, like preregistration, seem to be seen as procedures enabling researchers to explicate more procedural order in an effort to practically make others virtual witnesses (Lynch, 1985; Shapin & Schaffer, 1985; Shapin, 1984) and thereby reduce the, as Brenninkmeijer and colleagues (2025) put it, 'interpretative flexibility' (p.2) of the experimenters regress (see Chapter 7 for more on how preregistration links to flexibility in research practices).

Reproducibility and replication in the lab

While the most common understanding of the replication and reproducibility terminology seems to be the so-called Claerbout/Donoho/Peng convention (Barba, 2018; Matarese, 2022), the responses suggest that there is still quite some diversity and confusion in understanding as well as usage of the terminology. Indicating that in communities where such confusion and

[78] Original in German: „Ja, weil ich halt bestimmte Felder kenne, wo ich das Gefühl habe, Replikationen haben nicht gereicht. Also Kandidatengene, da gab es auch Replikationen, aber die positiven Replikationen ließen sich halt einfacher veröffentlichen als die Nonreplikationen, sodass die Meta-Analysen dann trotzdem zu falschen Schlüssen gelangt sind. […] Meine Nonreplikationen hatten da irgendeinen Impact, aber sicherlich noch höheren, dadurch dass sie präregistriert waren und Registered Reports wären davon wahrscheinlich noch besser gewesen.'

terminological proliferation are more prevalent, there can be cases where researchers using the same terms actually talk past each other by referring to different practices and epistemic functions that can be hidden behind the terms replication and reproducibility (Albertoni et al., 2023; Horbach et al., 2026; Ulpts & Schneider, 2024).

My observations and interviews suggest that 'real' or mere replication is not part of their normal research practice, due to a combination of perceived systematic forces incentivising and rewarding novelty as well as originality with their own interest to learn something novel and follow their own interests. Hence, 'real' or exact replication might just not be worth the necessary investment considering the intrinsic and extrinsic motivations of the researchers involved (Horbach et al., 2026; Latour & Woolgar, 1979; Mulkay & Gilbert, 1986; Peterson & Panofsky, 2021). Therefore, a practice that does appear to be part of their research routine is what I have termed 'partial replication', in which parts of previous research are taken (and adapted) to incorporate them as parts of their own research following their own interests and research questions. As previously stated, this is in accordance with findings from Mulkay and Gilbert (1986) of biochemists preferring the practice of replication through experimental variation over mere replication, Guttinger's (2018; 2019) insights about researchers in the life sciences practicing, what he called, micro-replications, and Peterson and Panofsky's (2021) findings concerning the preference for integrative replication among scientists. The preference for partial replication seems to be linked to the most prominent epistemic function behind the practice of replication and reproducibility voiced by the researchers, namely, to be able to build their own research on a stable foundation. Hence, the type of replication practice that can be observed seems to be less about the detection of mistakes, false-positives or even misconduct (Penders et al., 2020; Rensik & Shamoo, 2017), but more about integration and ensuring the ability to assure that something is useful for their own research so that they can further build on it due to a sense of trust and credibility that is provided by this form of (partial) replication. However, unlike in the case of Mulkay and Gilbert (1986), real (mere) replication was not rejected because it was just assumed to be possible anyway, quite the opposite, problems concerning the replicability of research were explicitly present in the researchers' responses (see e.g., Brenninkmeijer et al., 2025). These findings put into doubt how realistic or even useful current calls from Open Science reformers as well as research integrity proponents for more direct or exact replication across the sciences actually are for the researchers (see e.g., Peels & Bouter, 2023). Furthermore, differences in uses of terminology and hidden replications due to piloting together with the reservation of the term replication for only direct or exact replication might be some of the reasons why

metascientists' estimation of how many replications are practiced in science are usually so low and probably underestimate the amount of practiced replications as well as how many different forms replication can take in practice (see e.g., Clarke et al., 2024; Haig, 2022; Lubega et al., 2023; Penders et al., 2019; Ulpts & Schneider, 2024). These findings suggest that while the work on replication from classical SSK seems to be largely forgotten or ignored by research integrity scholars and Open Science reformers alike, they still might be relevant in discussions about the role, value and place of replication and reproducibility across the research landscape (see Brenninkmeijer et al., 2025). Hence, it can help with recognising and understanding areas of misalignment between current calls in the research integrity literature concerning the importance of direct replication and the perceptions and interests of the researchers, which can lead to potential frictions if these ideals about direct replication make it further into research integrity policy and evaluation (Labib et al., 2024; Valkenburg et al., 2021). Incorporation of classical SSK studies and their theoretical and methodological work in current science studies into the role and place of replication and reproducibility in science could provide intriguing future directions for research and a view into how the social dimension of these practices interacts with the epistemic concerns.

Chapter 9: Discussion and conclusion

This dissertation was prompted by the observation that research integrity has not only become a prominent category in science policy, governance, reform discourse and investigations of science (Aubert Bonn & Pinxten, 2019; Carcausto-Calla et al., 2023; Field et al., 2024; Hesselmann & Reinhart, 2024; NASEM, 2017; Tang, 2025b; Vendé et al., 2025; Zuckerman, 2020), but also a subject of contestation, because its relation to the researchers' own context dependent diverse and local understandings remains uncertain, with some studies indicating tensions between researchers' lived experiences and official notions of research integrity. Therefore, in this dissertation I set out to provide an emic perspective on the 'policy object' of research integrity (Davies & Lindvig, 2021; Sin, 2014), from within research laboratories (Pols et al., 2024).

To that end, I aimed to address the overarching research question of how researchers understand and enact research integrity or 'good research' in their everyday activities. This was divided into two sub-questions: 1) How do researchers talk with each other about and negotiate values and practices that have been linked to research integrity in the laboratory settings? 2) What motivations, concerns, and practical conditions shape these understandings and practices? Therefore, based on a constructivist approach, inspired by classical lab studies (Knorr-Cetina, 1981, 1995, 1999; Latour & Woolgar, 1979; Lynch, 1985; Traweek, 1988), and especially their insights on the role of regular meetings as sites of discourse about their processes of knowledge production, justifications and dissemination (Amann & Knorr-Cetina, 1989; Laudel, 2024; Müller, 2025; Müller et al., 2024; Philipps & Paruschke, 2024), I entered researchers' workplaces to conduct an ethnography of six scientific laboratories in Germany, the UK, and the Netherlands (Lynch, 2024; Stephens & Lewis, 2017). In an attempt to access their 'talking about science' and 'talking science', or 'integrity talk' and 'shop talk' respectively (Amann & Knorr-Cetina, 1989; Davies, 2019; Lynch, 1982, 1985; Philipps & Paruschke, 2024), I employed a combination of field observations during regular meetings, like lab meetings, journal clubs, and supervision meetings (Laudel, 2024; Müller, 2025), followed by 30 semi-structured interviews with the researchers at the end of each stay. To understand the scholars talking of 'good research' or 'good science' I framed their responses, interactions, explanations and justifications as forms of credibility management within their epistemic community as part of a larger transepistemic arena (Knorr-Cetina, 1982; Latour & Woolgar, 1979; Shapin, 1995a), with practices that serve the attainment and assessment of

credibility as a type of credibility work (Daston, 1995; Geiger, 2021; Shapin & Schaffer, 1985).

This dissertation provides three main contributions. First it complements the existing literature and empirical research on research integrity, which is mostly based on etic approaches that capture research integrity and 'good' or 'bad science' from the outside (see e.g., Allum et al., 2023; Gopalakrishna et al., 2022; Labib et al., 2021) , while at the same time adding to the already existing attempts at capturing an emic perspective of research integrity from inside the respective research communities or laboratories (see e.g., Brenninkmeijer & Pols, 2026; Penders et al., 2009; Pols et al., 2024; van Drimmelen et al., 2024, 2025). As such it further provides input for the elaboration and investigation of potential disconnects and misalignments of research on research integrity, research policy, research evaluation, and current science reform movements, with the realities, values, practices and preferences of researchers within their specific epistemic communities and context dependent conditions of knowledge production (see e.g., Brenninkmeijer & Pols, 2026; Labib et al., 2024; Nelson, 2021; Peterson, 2021; Pols et al., 2024; Robinson-Garcia et al., 2023; Valkenburg et al., 2021; Vincent & Hooper, 2025). This is outlined based on a combination of the researchers' personal reflections in conversations with each other (talking science) and with me (talking about science) concerning what constitutes 'good science' and its relation to the ethicality of scientific conduct, with the researchers' accounts of their practices of replication, reproducibility, and preregistration as well as the motivations behind them (Gläser, 2024; Latour & Woolgar, 1979).

The second contribution is the added insight to the still unclear relation of Open Science to diverse notions of 'good science' and research integrity (Haven et al., 2022). Due to the researchers' discourse about practices like preregistration and replication during their meetings, but also because of the active involvement of one of the research groups in issues surrounding Open Science and metascience, the material offers an intricate mixture of alignment and misalignment between Open Science and the researchers' conceptions of 'good research' as well as the conditions of their specific epistemic contexts. This can be put in relation to the existing literature on the relation of researchers' community specific values and preferences and their experienced realities within their environments with Open Science practices and policies (Cole et al., 2024; Davin et al., 2026; Guttinger, 2020; Leonelli, 2018, 2022, 2023; Lilja, 2021; Ma, 2025; Manco, 2024, 2025; Penders et al., 2019, 2020; Ulpts & Schneider, 2025).

The third contribution is the exploration of how classical science studies approaches, like lab ethnographies and theories about credibility management and research quality from the social studies of science can provide a

fruitful ground for investigation and comprehension of research integrity related concerns (Aradau & Huysmans, 2019; Latour & Woolgar, 1979; Shapin, 1995a), especially in the context of atheoretical tendencies in metascience and research integrity alike (Lambert & Degn, 2026; Peterson & Panofsky, 2023).

Good research from within: the emic perspective on research integrity

The main finding is that researchers do not recognise or use 'research integrity' as a familiar native category (Davies, 2019). However, this lack of familiarity itself is not necessarily a sign of disconnect or even misalignment between official notions of 'good science' and the researchers' understandings of 'good research'. Although, it becomes clear that following guidelines, teaching, and procedural ethics does not sufficiently explain what makes research good within local sites of knowledge production. The understandings and perceptions 'from within' are the result of their context specific negotiations between their specific aspirations of morality, credibility and responsibility on the one side, and the practical conditions at various levels under which their research is conducted, on the other (Valkenburg et al., 2021). However, the first theme from the fieldwork was that researchers were unfamiliar with the concept or term 'research integrity'. Furthermore, when trying to explicate what it might mean to them, they struggled to use it. Therefore, the themes presented are better seen as diffuse and somewhat messy associations rather than structured definitions. The themes that emerged in this free association exercise are: The morality and personality of the scholar, taking responsibility and ownership for ones work, honesty, humility and critical reflexivity, the role of the research system, the researcher in a group of researchers, the integrity of the research, misconduct: fabrication, falsification, and plagiarism, methodological rigour, hypothesis testing, research ethics ('do no harm'), prevention of bias, societal relevance and practical utility, and the relationship between research integrity and Open Science. Lastly, a theme that emerges from the reflections of the researchers across chapters is that there is perception of a general lack of research integrity in science.

The researchers' discourse demonstrates an intricate and at times confusing or contradictory and uncertain mixture of values or aspirations and responsibilities between stakeholders in the science system, like journals, funders, institutes, and the researchers themselves. While it remains unclear who is responsible for which component of research integrity, it emerges an image that makes clear that one cannot properly discuss responsibility for research integrity in isolation without acknowledging the interdependencies between these different stakeholders and types of actors involved. Although, a crucial

role is clearly allocated to the researchers in carrying the responsibility that the research is done well, and as previously indicated, even in hiring, funding and publishing decisions researchers are usually involved in the evaluation processes (Schneider et al., 2021).

While Open Science also has other motivations, like enabling and accelerating science's ability to feed into business and innovation (Fernández Pinto, 2020; Lund, 2025), it is usually presented as serving and ensuring research integrity (Bouter, 2024; Haven et al., 2022) or that Open Science and research integrity as well as research ethics can be mutually reinforcing (Lindemann & Häberlein, 2023). Researchers generally describe Open Science as a means to the end of making research accessible for use and evaluation. Thus, it is presented as one way to achieve the more general epistemic aim of 'Nachvollziehbarkeit' and 'Nachprüfbarkeit' (Ulpts & Schneider, 2025). The researchers indicate that Open Science and research integrity are not the same, or necessarily mutually reinforcing, as, according to the scholars, Open Science practices have to be implemented properly to work, they should not be applied like a checklist, some openly available material will never be accessed or checked, and a hyperfocus on one thing, like data sharing and statistical power can lead to detrimental effects by practically excluding certain types of research or lowering data quality (see e.g., Malich & Rehmann-Sutter, 2022).

Hence, similar to research ethics, Open Science, due to its multiplicity and differences in relevance and feasibility of specific Open Science practices across the diverse research landscape, further demonstrates the complexity, diversity and confused nature of research integrity, instead of disentangling it (Davin et al., 2026; Fecher & Friesike, 2014; Leonelli, 2023). Therefore, certain disconnects or misalignments between policy and institutional regulation in matters of research integrity, research ethics, and Open Science might be inevitable. On the one hand, regulation and policy require some form of standardisation, uniformity and simplicity to work (Vincent & Hooper, 2025), for instance in bureaucratic settings. On the other hand, there is diversity and context-dependency concerning not just the relevance and feasibility of different types of ethicality, openness, and integrity, but also regarding which forms they can take within local site of knowledge production.

Concerning the conceptual ambiguity of research integrity and its relation to research ethics, the researchers' inside perspective does not resolve said ambiguity, but actually demonstrates it, with some voicing frustration or talking about navigating around procedural ethics requirements (De Vries et al., 2006; Pickersgill, 2012; van Drimmelen et al., 2025) to be done with such a 'box ticking exercise' (Elizondo & Kaltenbrunner, 2024, p. 9), while it is also described as a component of 'good research', when ethics protocols are followed as they were approved by an ethics committee. The described co-

production and duality of care between ethical considerations for the participants' well-being and the goal of obtaining clean and reliable data (Brenninkmeijer & Pols, 2026; Mickel et al., 2023; Pickersgill, 2012), further conflates research ethics and research integrity as co-constitutive for research quality and 'good research'. Similarly, the ethicality of conduct within a research group as part of the lab climate, was described as a precondition for feeling comfortable, enabling some of the aspects they see as crucial for 'good research', like being honest about errors as well as reflexive and critical towards oneself and the group (Field et al., 2024; Valkenburg et al., 2021). However, the researchers also put a limit on how far the lab climate can alleviate stress and other aspects that they see as negatively affecting them and their research, because the lab is still situated in and dependent on the larger science system that is largely described as working against research integrity. They also describe how care for a good lab climate can conflict with notions of goo authorship practices.

Therefore, I would argue that since research integrity and research ethics are inherently intertwined conceptually and in practice (Kolstoe & Pugh, 2024; Pickersgill, 2012), especially in fields where the research is done with humans or animals, it would be negligent to not address the aspects of research ethics in discussions or conceptualisations of research integrity. Furthermore, for many the moral character of the researcher and the resulting conduct are paramount components of research integrity (see e.g., Daston, 1995; Shaw, 2019; Shaw & Satalkar, 2018; Zhaksylyk et al., 2023), making research integrity an inherently ethical subject. In the literature research ethics has also been described as the regulatory component of guidelines and research integrity as the behaviour of the researchers that is in accordance with such ethical guidelines (see e.g., Muthanna et al., 2024). Moreover, because research integrity is alien for researchers considering and talking about research ethics might provide another entry point to investigate and talk about research integrity related issues of 'good research' (Mickel et al., 2023). Therefore, my observations in the lab and the researchers' reflections on this link between 'good research' and ethics enacted as building rapport with participants and creating a save lab climate, instead of clearly separating them, conceptually and practically brings research ethics, research integrity and research quality closer together (Helgesson & Bülow, 2023; Kolstoe & Pugh, 2024).

Engaging with how they talked about specific practices linked to research integrity, like reproducibility, replication and preregistration, provides a window into how they enact 'good research' as well as reflect on such practices in the context of the conditions set by their local sites of knowledge production (Valkenburg et al., 2021). Such practice bound entry points to talking about

research integrity can further be complementary to the more abstract direct engagement with what constitute 'good research' or research integrity, especially since researchers struggle to explicate and define its nature when directly confronted with the concept or term research integrity (Hangel & Schickore, 2017). Their conversations and reflections of practices surrounding reproducibility, replication, and preregistration, demonstrate that they are generally committed to them as components of 'good research', due to their utility as tools in the management of the credibility of their research and themselves within their research community (Latour & Woolgar, 1979; Shapin, 1995a; Shapin & Schaffer, 1985). However, at the same time they talk about having to enact such practices in the context of limitations set by the nature of their research, the conditions set by the institute and wider science system, as well the workload associated with them.

Although organised scepticism is an old and idealised concept describing the self-organising phenomenon of inter-scholar criticism (Merton, 1942), its nature seems to have changed a bit over the last 60 years (Hosseini et al., 2024). In the past the organisation of scepticism was largely done by or in the context of scientific societies, who were responsible for the journals and scholarly communication. Now, in the age of Open Science and the internet this process is accelerated and more accessible (Hendriks & Reinhart, 2020; Meskus et al., 2018).

A component of ensuring research integrity conveyed by the researchers seems to be a certain anticipation of scrutiny, which I have termed a culture of suspicion, which becomes especially salient in the discourse on preregistration. Relatedly, in the context of sleuthing and online discourse about misconduct a similar phrasing turned up, this policing in an age of openness and accessibility has been called a 'vigilante culture' (Pulverer, 2014, p. 1286; also see Horbach & Valinciute, 2026). Hence, next to the more trust-based interdependence that the scholars in my fieldwork also view as part of research integrity or even precondition for collective knowledge production based on trust, there seems to be at the same time a culture of distrust in current academia as well as in the researchers' understanding of research integrity. Hence, researchers' conceptions of research integrity seem to be characterised by certain ambivalence or push and pull between trust and distrust, with the current system and research integrity community, apparently leaning more towards distrust, in the context of crisis narratives (Bartscherer & Reinhart, 2025; Penders, 2024; Peterson & Panofsky, 2023). Such a culture of distrust, monitoring, sleuthing and vigilantism could also contribute to transparency as a form of surveillance (Nguyen, 2022). Future research could be historical comparisons of scholarly discourse and criticism over various time intervals to see how the phenomenon of organised scepticism changes with time and

different mediums as well as organisations of communication dynamics. In a next step this could then be further explored by investigating how that influences research integrity and the role that critique and critical reflection play over time, as well as how they are received by scholars within the academic community (see e.g., 'Tone Debate' Derksen & Field, 2022, p. 172; Penders, 2022).

To a certain degree the employment of preregistration, as a reporting practice that aids researchers in persuading their research community of the credibility of their research, by making others virtual witnesses (Shapin, 1984; Shapin & Schaffer, 1985; Meskus et al., 2018), seems to be motivated by such a perception of a culture of suspicion. Such suspicion as a motivation behind the practice of preregistration and the aspect of preventing bias, as a component of 'good research', seem to be part of a moral economy of noninterventionist objectivity (Daston, 1995; Daston & Galison, 1992), or as Daston and Galison put it: 'all epistemology begins in fear' (Daston & Galison. 2007, p. 372). Considering the classical lab studies conceptualising the production of matters of facts and the reporting in publications as processes of making order out of chaos (Latour & Woolgar, 1979; Lynch, 1985), preregistration's utility as a credibility management tool can also be linked to its perceived capacity to create an impression of order, for instance, as the time-stamped protocol outlining the sequence of events and separating exploration from confirmation.

Replication and reproducibility are among the themes that occupy a centre stage in the current research integrity discourse, probably propelled by a moral panic associated with the crisis narrative underlying a lot of Open Science and metascience diagnoses and recommendations (Bartscherer & Reinhart, 2025; Hesselmann & Reinhart, 2024; Penders, 2024; Penders et al., 2020; Peterson & Panofsky, 2023). In his little play of a conversation between scientists and sociologists of science, Mulkay (1984) points towards the irony of replicating SSK studies on replication and making arguments about added credibility based on the value of such successful replications, considering their own arguments concerning the experimenters' regress, the aspect of social negotiation or non-epistemic factors in replication practice and appraisal, and the influence of tacit knowledge (see e.g., Ashmore, 1988; Brenninkmeijer et al., 2025; Collins, 1975; 1985; Travis, 1981). I find myself in exactly that irony, of apparently replicating sociology of science (scientific knowledge) findings on replication (Brenninkmeijer et al., 2025; Collins, 1975, 1981, 1985; Mulkay & Gilbert, 1986; Guttinger, 2019; Perterson & Panofsky, 2021; Travis, 1981). Such irony provides an intriguing invitation to be reflexive and situate and relativise one's own social practice of replication or follow something like Collins' methodological relativism and only suspend realism for the duration of

investigation to prevent getting stuck in a tautology, but turn the realism back on once the investigation is done (Franklin & Collins, 2016).

My findings are largely congruent with Mulkay and Gilbert's (1986) as well as Peterson and Panofsky's (2021) coining of integrative replication and Guttinger's (2019) notion of micro-replication. Direct replications are not described as particularly valuable or worth the necessary investment considering the expected benefit for them and their research (Latour & Woolgar, 1979). Therefore, they are also not part of the experiences the researchers talk about, with one or two exceptions from scholars who are themselves involved in metascience. However, what they do talk about, as their practice, is 'partial replication' where they take parts of previous research and see how they can fit it into their research workflow to follow their own research interests (Guttinger, 2019; Mulkay & Gilbert, 1986; Peterson & Panofsky, 2021). As such, unlike current presentations of replication's and reproducibility's utility for research integrity, the practice is said to be less employed due to its believed value in diagnosing bad conduct like questionable research practices and misconduct or allocating responsibility (Penders et al., 2020; Resnik & Shamoo, 2017), but rather to partially build one's own research on a stable foundation, provided by the added credibility of the replicated part, while still being able to incorporate something novel (Guttinger, 2019; Mulkay & Gilbert, 1986; Peterson & Panofsky, 2021).

With lab studies from science studies over credibility to research integrity

If the aim of research integrity, whether in the form of policy or the community dedicated to research on research integrity is to ensure and improve proper conduct (see e.g., Aubert Bonn & Pinxten, 2019; Carcausto-Calla et al., 2023; Nguyen & Tuamusk, 2024; Soehartono & Khor, 2022; Tang, 2025b; Vendé et al., 2025), then a fundamental prerequisite for appropriate implementations as well as relevant and feasible interventions across different epistemic contexts is to actually understand how and under which conditions different types of research are really enacted (Nelson, 2021; Peterson, 2021). Science studies, and especially the social studies of science, have a long tradition of investigating research in various ways (Pickering, 1992; Shapin, 1995b), and demonstrating epistemic diversity or the disunity of the sciences (Galison & Stump, 1996; Knorr-Cetina, 1999; Laudel, 2024). In this dissertation I mainly drew from theories, insights and methodologies that emerged in the history of science studies in the context of the sociology of scientific knowledge and cultural approaches to studying the sciences (see e.g., Knorr-Cetina, 1995; Shapin, 1995b). In other words, I mainly drew from work of the second wave of science

studies, while recent research integrity communities, metascientists, and reform agendas seem to revert back to the first wave of science studies (Collins & Evans, 2002). Ethnographies of scientific practices and discourse within their local sites of knowledge production, like laboratories, provide a unique form of access for the investigation of how researchers negotiate their conduct in interaction with the conditions set by their specific context (Lynch, 2024; Stephens & Lewis, 2017).

Considering that research integrity, while being an ill-defined and elusive umbrella concept (Field et al., 2024; Kolstoe & Pugh, 2024; Valkenburg et al., 2021), is in the literature often linked to issues of research credibility (Kretser et al., 2019; Naddaf, 2025; Pols et al., 2024), the theoretical framing, based on social studies of science work on credibility, the transepistemic arena, the credibility cycle, and researchers' motivations to engage with specific practices (see e.g., Knorr-Cetina, 1982; Latour & Woolgar, 1979; Shapin, 1995a; Smith, 2014), aided the analysis as well as in understanding how researchers reflections and practices relate to their notions of 'good research' and ultimately contribute to an emic perspective on research integrity. The framing of research integrity within a transepistemic arena further allows to recognise research integrity as a 'policy object' (Davies & Lindvig, 2021; Sin, 2014), while situating the credibility cycle within that arena allowed for conceptualising a process of negotiation and translation between official notions of research integrity and the researchers' own notions of 'good research' in practice. Such and other pre-existing literature on the nature of the sciences their interdependencies with other stakeholders like publishers, funders, businesses, think tanks, and their role in society provide a crucial source for insight and inspiration, with an assortment of theories, perspectives and methodologies.

As the response of one of the researchers, that research integrity captures how laypeople think science is conducted, suggests, research integrity might also be a case of the public understanding of science finding its way back into science as an image of science that is supposed to guide scientific practice, but seems, therefore, to be somewhat estranged from the actual reality of research conduct. Thus, while some argue that part of research integrity is or should be responsible science communication (Ciubotariu & Bosch, 2022), the explication, and especially codification, of 'good science' as research integrity in research policy and guidelines could also be treated as an act of science communication itself, where scientists are themselves a crucial audience. We even observe some of the same issues emerging in the research integrity and the science communication literature, especially in the context of declarations of a crisis or post truth state of science. Such domain crossing themes are, for instance, how to engage with pluralism, uncertainty, and the value-ladenness of research (see e.g., Ambrosj et al., 2023; Kaldewey, 2025; Penders et al., 2009).

As such research integrity might be the consequence of a curious looping effect (Hacking, 2007), taking place in the context of current scandalisations and a moral panic about research misconduct and questionable research practices (Bartscherer & Reinhart, 2025; Hesselmann & Reinhart, 2024; Horbach et al., 2024; Penders, 2024; Peterson & Panofsky, 2023). Something else has to be considered about notions of 'good science' or 'good research' that can be found in guidelines, policies or interview situations: When researchers are asked to explicate what is often tacit and embedded in lived practice, and to translate their understandings of 'good research' to wider audiences and outsiders, their accounts may also become ideological representations of 'good research' — whether deliberately, as boundary work, or unintentionally. (Gieryn, 1983; Gläser, 2024; Gulbrandsen, 2000; Mulkay, 1976; Vincent & Hooper, 2025).[79]

Research integrity being a collection of ideals or an ideology, itself does not have to be of concern per se. However, if these notions of 'good science' are used as standards for the reward and punishment of researchers, instead of ideals that they aspire towards, research integrity instead of being a 'useful fiction' could become a new form of scientism. As classical lab ethnographies suggest, science is not necessarily conducted in a way as laypeople think it is; it is more diverse, uncertain, messier and contingent on social, personal, political, and cultural circumstances of the local contexts (see e.g., Collins, 1985; Knorr-Cetina, 1995; 1999; Latour & Woolgar, 1979; Lynch, 1885; van Drimmelen et al., 2024, 2025). Therefore, research integrity as a concept and the codification of notions of 'good science' (Vincent & Hooper, 2025), for instance, in guidelines and policy, might carry the danger that: 'By making the ideal picture of science more idealistic, it made actual research look worse' (Schniedermann, 2025, p.21). Which can introduce a vicious cycle, where the measures to improve or ensure 'good science' actually create the impression that the state of science is worsening and, therefore, requiring further countermeasures like guidelines, incentives, teaching, and regulation.

I propose that science studies literature on science communication and historical episodes where an idealised understanding of science was confronted with issues like the messiness, diversity, uncertainty and context-dependency of science in action (e.g., the Science Wars and Climategate) can provide fruitful grounds to explore how to ensure that conceptualisations and codifications of research integrity do not slip into scientistic notions of science and confuse aspirations for 'good research' with standards (Garud et al., 2014; Lynch, 2020).

[79] Ochsner (2022) even treats research quality as a latent construct. Which begs the question whether emic notions of good research are generally rather latent constructs instead of directly observable ones.

Limitations

In the dissertation the interview material is more visible and represented than the fieldnotes from the observations during the meetings. This can have, at least, two reasons: 1) expressions of research integrity and related issues in verbal interactions between researchers could have been harder to identify and observe, because research integrity is not part of their native vocabulary. As such the more inductive than deductive procedure of taking field notes complicated the process of capturing 'integrity talk' (Davies, 2019), 2) research integrity, similar to research quality, might be a rather tacit concept or latent construct (Gulbrandsen, 2000; Ochsner, 2022), which practically hides research integrity in the laboratories behind daily practice and routine. Therefore, often longer durations of field observations for immersion are required to properly make the invisible parts of the researchers' enacted components of 'good research' visible (van Drimmelen et al., 2024).

The higher reliance on interview data further highlights the crucial effect that the interview situation can have on the scientists' responses (Gibert & Mulkay, 1984). While, I based the interviews on the themes that emerged in the observable discourse between researchers during the field observations, thereby trying to anchor their reflections in their actual practices and experiences, the danger remains that what I captured was more an ideology of science due to the process of translating insider notions to an outsider, attempts at explicating the tacit, and social desirability of the researchers wanting to preserve a certain image of science and themselves (Gieryn, 1983; Gläser, 2024; Hangel & Schickore, 2017; Mulkay, 1976; Tavory, 2020). Moreover, the specific conceptualisations and perspectives provided by the scholars, could to a non-negligible degree stem from my framing of research integrity as 'good research' or 'good science', as a way to overcome the interviewees struggle with the term and concept research integrity.

Furthermore, because I mostly only observed conversations between scientist and conducted interviews, this is mostly just a representation of what they say and are willing to share with me. Therefore, as research into the normative dissonance of scholars demonstrates, what researchers say they do or would do, does not have to be congruent with their actual conduct (see e.g., Anderson et al., 2007). Hence, inferences to actual research practices based on this material requires caution and humility. For that purpose, longer ethnographic fieldwork might be necessary, where the ethnographer can further immerse him- or herself into the research enactment situations to decrease the influence of social desirability and disturbances caused by the ethnographer as well as have a greater chance at identifying and describing tacit components of knowledge production and justification.

Moreover, as my problems with getting access to research labs, as described in the method chapter, already illustrated, there is of course the risk of a certain combination of selection and response bias concerning the field sites and researchers I investigated. This potentially limits the transferability, generalisability and representativeness of my findings, interpretations, and conclusions, not just beyond the boundaries of the cognitive, social and behavioural sciences, but even within them (Yadav, 2022). Furthermore, the western European focus in the project most likely limits the applicability and transferability to non-western European countries where the enactment situation might be different. Also, the perception of Open Science and how it relates to research integrity will most likely differ, because some of the researchers in my study are actively involved in Open Science, and the biomedical, behavioural, cognitive and social sciences are currently the epicentre of Open Science discourse and interventions.

Additionally, since the research I investigated relies on human (or animal) participation, the ethical components of research in disciplines lacking such participants might differ, as might the relationship between research integrity and research ethics. Further research is needed to investigate perceptions of research integrity and its relation to Open Science as well as research ethics, in for instance, engineering, the humanities, and the natural sciences (see e.g., Davin et al., 2026; Manco, 2025). This would also allow for the elaboration and exploration of different types of practices, other than preregistration, reproducibility and replication, which have been linked to different notions of 'good research' or research integrity across research approaches, disciplines, and fields.

Conclusion

All in all, Research integrity is not absent from everyday research simply because researchers do not use the term or are unfamiliar with codes of conduct (Davies, 2019). Furthermore, previous claims of misalignment between guidelines or policy and research realities might also be the consequence of mistaking the map for the territory (Vincent & Hooper, 2025), because it seems simply unrealistic to expect research integrity policy and guidelines do neatly map onto the various differing research realities across, disciplines, institutes and grographic locations. As such, the question seems to be less about misalignment between official notions of research integrity and researchers' understandings of 'good research', but rather about application of policy and requirements as well as incentives, to prevent cases of policy alienation or epistemic injustices (Fricker, 2007; Lilja, 2021; Penders et al., 2019; Tummers et al., 2012). We find the researchers' own notions of 'good research' situated within the conditions of the local sites of knowledge production, as well as the

wider science system (Pols et al., 2024). These concerns revolve around issues like how to make the research credible, be a moral scholar, allocate responsibilities, navigate the interdependencies with other researchers, and, as the reflections on the practices of preregistration, replication and reproducibility illustrate, find a balance between costs and benefits regarding the necessary investment for certain practices against their expected utility in the management of their credibility and recognition in the transepistemic arena of the sciences (Knorr-Cetina, 1982; Shapin, 1995a).

Considering the influx of Open Science narratives and interventions in the research integrity discourse and repertoire of practices, the characterisation of Open Science as a means to the end of 'good research', highlights the caution to not fall victim to Campbell's law, let Open Science become a superficial veneer of credibility, or incentivise and require specific Open Science practices irrespective of considerations about their relevance and feasibility for the research and context at hand (Devezer et al., 2021; Devezer & Penders, 2023; Leonelli, 2023; Ulpts & Schneider, 2025).

Furthermore, the inherent connection between ethicality of conduct in interaction with participants as well as one's colleagues and the ability to practice 'good research', for example, in the sense of being able to be reflexive or getting clean and reliable data, makes research ethics not something separate, but a fundamental component of research integrity (as 'good research') (Mickel et al., 2023; Pickersgill, 2012). On a procedural and bureaucratic level, it might be possible to cleanly distinguish between research ethics and research integrity, but in practice ethicality and integrity seems intertwined. Thus, whether routine, 'good' or 'bad conduct', research activity is inherently value-laden and fundamentally characterised by dynamics of community and culture specific forms of morality (Daston, 1995).

Summary

In this monograph I present the findings of an ethnographic study on researchers' perceptions, understandings, and experiences related to research integrity, conducted in cognitive science, psychology, and neuroimaging labs in Germany, the UK, and the Netherlands. Research integrity, simply put 'good science', has been a hotly debated topic within the last 20 years. This increased attention on how science is doing, and the proper conduct of researchers stems from widely circulated claims about a so-called 'replication crisis' or 'credibility crisis' in science. This crisis narrative comprises of issues like: a widespread lack of replicable research findings, problematic numbers of (high profile) scientific misconduct cases (like Diederik Stapel), a lack of transparency and openness in science, alarming rates of questionable research practices (research conduct that deviates from agreed upon community standards, norms and rules, but does not rise to the level of fabrication, falsification or plagiarism), an increase in numbers of retracted publications (publications are marked as problematic and not trustworthy), high-output predatory journals producing publications of low quality, and this all in a the context of an incentive structure in science that rewards scholars for quantity over quality. Therefore, activism and scholarship have mobilised to address these issues by investigating the ins and outs of scientific research, as well as to propose reforms regarding how to secure and foster research integrity.

However, the current discourse about research integrity is heavily confused. On the one hand there is an influx of new communities like metascience and the Open Science movement using research integrity terminology. While, on the other hand, research integrity scholars are addressing metascience and Open Science reforms. This is further complicated by the similarly confused usage of so many other closely related concepts, such as responsible conduct of research, research ethics, research quality, research governance, scientific integrity, personal integrity, as well as, just simply, 'good science'. Furthermore, previous research indicates a disconnect and areas of misalignment between researchers and policy as well as research integrity scholarship alike, when it comes to what actually constitutes research integrity. Likewise, there is an insufficient knowledge base about how exactly researchers relate to the 'policy object' research integrity and how they would describe what this concept is trying to capture. Hence, I set out on a voyage to see what research integrity as 'good research' looks like and means 'from within' the laboratory.

I illustrate that the term research integrity is not familiar to the researchers, and they do not know the underlying concept. However, once they start to

think about what this term might mean, in the context of their research, quite the complex and multifaceted picture emerges. Subsequently, I explore the relationship between research integrity and research ethics. It turns out that they are hardly separable in practice, and according to the researchers' conceptualisations research integrity is co-produced by epistemic and ethical considerations of knowledge production. Due to their prominence in the current research integrity discourse, especially in the social, cognitive and behavioural sciences, and them being part of large-scale science reforms, afterwards I focus on the practices of preregistration, replication, and reproducibility. The reporting practice preregistration is presented as an intricate mixture of providing a way to communicate and assess credibility on the one hand, and a practice adding uncertainty and extra workload that sometimes feels like punishment, on the other. Furthermore, it turns out that, in line with classical social studies of science, researchers do not talk about conducting direct or exact replications, what they do is take parts of previous research and see how they can incorporate them in their own research.

My findings provide an insightful inside perspective of researchers' understandings and perceptions of research integrity. Moreover, in my report an intricate and complex mixture of alignment and misalignment between researchers' realities and research integrity scholarship and policy emerges. In the end I provide a succinct summary of the findings and an outlook of why the insights, history, methodologies and theories of science studies provide a fruitful ground for future directions for research integrity policy and research alike.

Dansk resumé

I denne monografi præsenterer jeg resultaterne af en etnografisk undersøgelse af forskeres opfattelser, forståelse og erfaringer i relation til forskningsintegritet, som er gennemført på laboratorier inden for kognitionsvidenskab, psykologi og neuroimaging i Tyskland, Storbritannien og Holland. Forskningsintegritet, kort sagt 'god videnskab', har været et meget omdiskuteret emne inden for de sidste 20 år. Denne øgede opmærksomhed på, hvordan videnskaben klarer sig, og forskeres korrekte adfærd stammer fra vidt udbredte påstande om en såkaldt 'replikationskrise' eller 'troværdighedskrise' inden for videnskaben. Denne krisefortælling omfatter emner som: en udbredt mangel på replikerbare forskningsresultater, et problematisk antal (højt profilerede) tilfælde af videnskabelig uredelighed (som Diederik Stapel), en mangel på gennemsigtighed og åbenhed i videnskaben, alarmerende tal for tvivlsomme forskningspraksisser (forskningsadfærd, der afviger fra aftalte standarder, normer og regler i forskersamfundet, men som ikke når op på niveauet for fabrikation, forfalskning eller plagiering), en stigning i antallet af tilbagekaldte publikationer (publikationer, der er markeret som problematiske og upålidelige), højproduktive rovtidsskrifter, der producerer publikationer af lav kvalitet, og alt dette i en kontekst med en incitamentsstruktur inden for videnskaben, der belønner forskere for kvantitet frem for kvalitet. Derfor er aktivister og forskere gået sammen om at tackle disse problemer ved at undersøge de mange facetter af videnskabelig forskning samt foreslå reformer, der skal sikre og fremme forskningsintegriteten.

Den aktuelle debat om forskningsintegritet er imidlertid præget af stor forvirring. På den ene side ser vi en bølge af nye fællesskaber som metavidenskab og Open Science-bevægelsen, der anvender terminologi fra forskningsintegriteten. På den anden side beskæftiger forskere inden for forskningsintegritet sig med reformer inden for metavidenskab og Open Science. Dette kompliceres yderligere af den tilsvarende forvirrede brug af så mange andre nært beslægtede begreber, såsom ansvarlig forskningsadfærd, forskningsetik, forskningskvalitet, forskningsstyring, videnskabelig integritet, personlig integritet samt ganske enkelt 'god videnskab'. Desuden tyder tidligere forskning på en kløft og uoverensstemmelser mellem forskere og politik såvel som forskning i forskningsintegritet, når det kommer til, hvad der egentlig udgør forskningsintegritet. Ligeledes er der en utilstrækkelig videnbase om, hvordan forskere præcist forholder sig til det 'politiske objekt' forskningsintegritet, og hvordan de ville beskrive, hvad dette begreb forsøger at indfange. Derfor begav jeg mig

ud på en rejse for at se, hvordan forskningsintegritet som ›god forskning‹ ser ud og betyder 'indefra' laboratoriet.

Jeg viser, at begrebet forskningsintegritet ikke er kendt blandt forskerne, og at de ikke kender det underliggende begreb. Men så snart de begynder at overveje, hvad dette begreb kan betyde i sammenhæng med deres forskning, tegner der sig et ganske komplekst og mangesidigt billede. Derefter undersøger jeg forholdet mellem forskningsintegritet og forskningsetik. Det viser sig, at de i praksis næppe kan adskilles, og ifølge forskernes opfattelse er forskningsintegritet et resultat af epistemiske og etiske overvejelser om videnproduktion. På grund af deres fremtrædende rolle i den aktuelle diskurs om forskningsintegritet, især inden for samfunds-, kognitions- og adfærdsvidenskaberne, og fordi de indgår i omfattende reformer af videnskaben, fokuserer jeg derefter på praksis inden for forhåndsregistrering, replikation og reproducerbarhed. Forhåndsregistrering præsenteres som en indviklet blanding af på den ene side at give en måde at kommunikere og vurdere troværdighed på, og på den anden side en praksis, der tilføjer usikkerhed og ekstra arbejdsbyrde, som nogle gange føles som en straf. Desuden viser det sig, at forskere, i tråd med klassiske samfundsvidenskabelige studier af videnskab, ikke taler om at gennemføre direkte eller nøjagtige replikationer; det, de gør, er at tage dele af tidligere forskning og se, hvordan de kan inkorporere dem i deres egen forskning.

Mine resultater giver et indsigtsfuldt indblik i forskernes forståelse af og opfattelse af forskningsintegritet. Desuden tegner der sig i min rapport et indviklet og komplekst billede af, hvorvidt der er overensstemmelse eller uoverensstemmelse mellem forskernes virkelighed og den akademiske litteratur og politikken omkring forskningsintegritet. Afslutningsvis giver jeg et kortfattet resumé af resultaterne samt et overblik over, hvorfor indsigterne, historien, metoderne og teorierne inden for videnskabsstudier udgør et frugtbart grundlag for den fremtidige retning inden for både politikken for forskningsintegritet og selve forskningen.

References

Abbott, A. (1999). German science admits to fraud. *Nature*, *398*(6730), 765–766. https://doi.org/10.1038/19673

Abbott, A. (2019). The science institutions hiring integrity inspectors to vet their papers. *Nature*, *575*(7783), 430–433. https://doi.org/10.1038/d41586-019-03529-w

Aczel, B. (2023). Why we need a 'psychology of science.' *Nature Human Behaviour*, *8*(1), 4–5. https://doi.org/10.1038/s41562-023-01786-4

Albertoni, R., Colantonio, S., Skrzypczyński, P., & Stefanowski, J. (2023). *Reproducibility of Machine Learning: Terminology, Recommendations and Open Issues* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.2302.12691

ALLEA - All European Academies. (2023). *The European Code of Conduct for Research Integrity*. ALLEA - All European Academies. https://doi.org/10.26356/ECoC

Allum, N., Reid, A., Bidoglia, M., Gaskell, G., Aubert-Bonn, N., Buljan, I., Fuglsang, S., Horbach, S., Kavouras, P., Marušić, A., Mejlgaard, N., Pizzolato, D., Roje, R., Tijdink, J., & Veltri, G. (2023). Researchers on research integrity: A survey of European and American researchers. *F1000Research*, *12*, 187. https://doi.org/10.12688/f1000research.128733.1

Amann, K., & Knorr Cetina, K. (1989). Thinking through talk: An ethnographic study of a molecular biology laboratory. *Knowledge and Society*, *8*, 3–26.

Ambrosj, J., Dierickx, K., & Desmond, H. (2023). The Value-Free Ideal of Science: A Useful Fiction? A Review of Non-epistemic Reasons for the Research Integrity Community. *Science and Engineering Ethics*, *29*(1), 1. https://doi.org/10.1007/s11948-022-00427-9

Amorim, M. M. R., & Bacchi, A. D. (2025). The crisis in scientific integrity and its implications. *Revista Brasileira de Saúde Materno Infantil*, *25*, e20240368. https://doi.org/10.1590/1806-9304202500000368-en

Anderson, M. S., Martinson, B. C., & De Vries, R. (2007). Normative Dissonance in Science: Results from a National Survey of U.S. Scientists. *Journal of Empirical Research on Human Research Ethics*, *2*(4), 3–14. https://doi.org/10.1525/jer.2007.2.4.3

Anderson, M. S., Ronning, E. A., Devries, R., & Martinson, B. C. (2010). Extending the Mertonian Norms: Scientists' Subscription to Norms of Research. *The Journal of higher education*, *81*(3), 366–393. https://doi.org/10.1353/jhe.0.0095

Anderson-González, M., & Gruson-Daniel, C. (2023). *Mapping contemporary "research on research" and "science studies": How new methods change the traditional academic landscape and inform public open science policies*. MetaArXiv. https://doi.org/10.31222/osf.io/p4mca

Angrist, J. D., & Pischke, J.-S. (2010). The Credibility Revolution in Empirical Economics: How Better Research Design is Taking the Con out of Econometrics. *Journal of Economic Perspectives*, *24*(2), 3–30. https://doi.org/10.1257/jep.24.2.3

Ankel-Peters, J., Fiala, N., & Neubauer, F. (2023). Do economists replicate? *Journal of Economic Behavior & Organization*, *212*, 219–232. https://doi.org/10.1016/j.jebo.2023.05.009

Announcement: Reducing our irreproducibility. (2013). *Nature*, *496*(7446), 398–398. https://doi.org/10.1038/496398a

Aradau, C., & Huysmans, J. (2019). Assembling credibility: Knowledge, method and critique in times of 'post-truth.' *Security Dialogue*, *50*(1), 40–58. https://doi.org/10.1177/0967010618788996

Armond, A. C. V., Cobey, K. D., & Moher, D. (2024). Research Integrity definitions and challenges. *Journal of Clinical Epidemiology*, *171*, 111367. https://doi.org/10.1016/j.jclinepi.2024.111367

Armond, A. C. V., & Kakuk, P. (2022). Perceptions of Research Integrity Climate in Hungarian Universities: Results from A Survey among Academic Researchers. *Science and Engineering Ethics*, *28*(4), 30. https://doi.org/10.1007/s11948-022-00382-5

Ashmore M. (1988). The life and opinions of a replication claim: Reflexivity and symmetry
in the sociology of scientific knowledge. In Woolgar S. (Ed.), *Knowledge and reflexivity: New frontiers in the sociology of knowledge* (pp. 125–153). Sage.

Aubert Bonn, N., & Pinxten, W. (2019). A Decade of Empirical Research on Research Integrity: What Have We (Not) Looked At? *Journal of Empirical Research on Human Research Ethics*, *14*(4), 338–352. https://doi.org/10.1177/1556264619858534

Aunger, R. (1995). On Ethnography: Storytelling or Science? *Current Anthropology*, *36*(1), 97–130. JSTOR.

Azoulay, P., Furman, J. L., Krieger, J. L., & Murray, F. E. (2012). *Retractions* (NBER Working Papers No. 18499). National Bureau of Economic Research, Inc. https://doi.org/None

Bacevic, J., & Muellerleile, C. (2018). The moral economy of open access. *European Journal of Social Theory*, *21*(2), 169–188. https://doi.org/10.1177/1368431017717368

Bak-Coleman, J., & Devezer, B. (2024). Claims about scientific rigour require rigour. *Nature Human Behaviour*, *8*(10), 1890–1891. https://doi.org/10.1038/s41562-024-01982-w

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature*, *533*(7604), 452–454. https://doi.org/10.1038/533452a

Barba, L. A. (2018). *Terminologies for Reproducible Research* (Version 1). arXiv. https://doi.org/10.48550/ARXIV.1802.03311

Bartscherer, S. F. (2026). Cultures of Trial and Error: The Narrative Side of (Open) Science. *Blog of Trial and Error*. https://doi.org/10.36850/448e9a73-75b2

Bartscherer, S. F., & Reinhart, M. (2025). *The (Non)Academic Community Forming around Replications: Mapping the International Open Science space via its Replication Initiatives*. https://doi.org/10.31235/osf.io/rbyt6_v2
Bazzoli, A. (2022). Open science and epistemic pluralism: A tale of many perils and some opportunities. *Industrial and Organizational Psychology*, *15*(4), 525–528. https://doi.org/10.1017/iop.2022.67
Becher, T. (1989). Academic Tribes and Territories: Intellectual Inquiry and the Cultures of Disciplines. Milton Keynes: SRHE/OUP.
Bechky, B. A. (2020). *Blood, Powder, and Residue: How Crime Labs Translate Evidence into Proof*. Princeton University Press. JSTOR. https://doi.org/10.2307/j.ctv13qfvpn
Bem, D. J. (2011). Feeling the future: Experimental evidence for anomalous retroactive influences on cognition and affect. *Journal of Personality and Social Psychology*, *100*(3), 407–425. https://doi.org/10.1037/a0021524
Benedictus, R., Miedema, F., & Ferguson, M. W. J. (2016). Fewer numbers, better science. *Nature*, *538*(7626), 453–455. https://doi.org/10.1038/538453a
Bergkvist, L. (2020). Preregistration as a way to limit questionable research practice in advertising research. *International Journal of Advertising*, *39*(7), 1172–1180. https://doi.org/10.1080/02650487.2020.1753441
Bezuidenhout, L., & Beeler, D. (2021). Dynamic Boundaries: Using Boundary Work to Rethink Scientific Virtues. In E. Ratti & T. A. Stapleford (Eds.), *Science, Technology, and Virtues* (1st ed., pp. 269–286). Oxford University PressNew York. https://doi.org/10.1093/oso/9780190081713.003.0015
Biagioli, M., & Lippman, A. (Eds.). (2020). *Gaming the Metrics: Misconduct and Manipulation in Academic Research*. The MIT Press. https://doi.org/10.7551/mitpress/11087.001.0001
Bimber, B., & Guston, D. H. (1995). Politics by the Same Means Government and Science in the United States. In S. Jasanoff, G. Markle, J. Peterson, & T. Pinch, *Handbook of Science and Technology Studies* (pp. 554–571). SAGE Publications, Inc. https://doi.org/10.4135/9781412990127.n24
Blatch-Jones, A. J., Lakin, K., & Thomas, S. (2024). A scoping review on what constitutes a good research culture. *F1000Research*, *13*, 324. https://doi.org/10.12688/f1000research.147599.3
Bolland, M. J., Grey, A., & Avenell, A. (2022). Citation of retracted publications: A challenging problem. *Accountability in Research*, *29*(1), 18–25. https://doi.org/10.1080/08989621.2021.1886933
Boltanski, L., & Thévenot, L. (2006). *On justification: Economies of worth*. (C. Porter, Trans.). Princeton University Press.
Bosch, X. (2010). Safeguarding good scientific practice in Europe. *EMBO Reports*, *11*(4), 252–257. https://doi.org/10.1038/embor.2010.32
Bourdieu, P. (1975). The specificity of the scientific field and the social conditions of the progress of reason. *Social Science Information*, *14*(6), 19–47. https://doi.org/10.1177/053901847501400602

Bouter, L. (2024). Why research integrity matters and how it can be improved. *Accountability in Research*, *31*(8), 1277–1286. https://doi.org/10.1080/08989621.2023.2189010

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

Bouter, L. M., & Riet, G. T. (2021). Replication Research Series-Paper 2: Empirical research must be replicated before its findings can be trusted. *Journal of Clinical Epidemiology*, *129*, 188–190. https://doi.org/10.1016/j.jclinepi.2020.09.032

Brandt, M. J., IJzerman, H., Dijksterhuis, A., Farach, F. J., Geller, J., Giner-Sorolla, R., Grange, J. A., Perugini, M., Spies, J. R., & van't Veer, A. (2014). The Replication Recipe: What makes for a convincing replication? *Journal of Experimental Social Psychology, 50,* 217–224. https://doi.org/10.1016/j.jesp.2013.10.005

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp0630a

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, *11*(4), 589–597. https://doi.org/10.1080/2159676X.2019.1628806

Braun, V., & Clarke, V. (2023). Toward good practice in thematic analysis: Avoiding common problems and be(com)ing a *knowing* researcher. *International Journal of Transgender Health*, *24*(1), 1–6. https://doi.org/10.1080/26895269.2022.2129597

Brenninkmeijer, J., Derksen, M., Meirmans, S., & Pols, J. (2025). Replication after Collins: An ethnography of current replication studies in psychology. *Social Studies of Science*, 03063127251382558. https://doi.org/10.1177/03063127251382558

Brenninkmeijer, J., Derksen, M., & Rietzschel, E. (2019). Informal Laboratory Practices in Psychology. *Collabra: Psychology*, *5*(1), 45. https://doi.org/10.1525/collabra.221

Brenninkmeijer, J., & Pols, J. (2026). Good chemistry and its interruptions: About caring scientists and uncooperative molecules. *Science as Culture*, *35*(1), 59–80. https://doi.org/10.1080/09505431.2025.2555180

Brodeur, A., Cook, N., Hartley, J., & Heyes, A. (2022). Do Pre-Registration and Pre-Analysis Plans Reduce p-Hacking and Publication Bias? *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.4180594

Brodeur, A., Cook, N. M., Hartley, J. S., & Heyes, A. (2024). Do Preregistration and Preanalysis Plans Reduce *p* -Hacking and Publication Bias? Evidence from 15,992 Test Statistics and Suggestions for Improvement. *Journal of Political Economy Microeconomics*, *2*(3), 527–561. https://doi.org/10.1086/730455

Buck, S. (2024). We Should Do More Direct Replications in Science. *Harvard Data Science Review*, *6*(3). https://doi.org/10.1162/99608f92.4eccc443

Cadwallader, L., Morton, L., & Hrynaszkiewicz, I. (2023, April 4). Trends in preprint, data, and code sharing, 2019-2022. *The Official PLOS Blog*. https://theplosblog.plos.org/2023/04/open-science-indicators/

Campbell, D. T. (1979). Assessing the impact of planned social change. *Evaluation and Program Planning*, *2*(1), 67–90. https://doi.org/10.1016/0149-7189(79)90048-X

Carcausto-Calla, W., Zapata, N. A., Cueva, F. E. I., Morales, S. A. G., & Rios, A. R. (2023). Mapping of Empirical Studies on Research Integrity in University Institutions. *Academic Journal of Interdisciplinary Studies*, *12*(5), 9. https://doi.org/10.36941/ajis-2023-0122

Carter, S. M., & Little, M. (2007). Justifying Knowledge, Justifying Method, Taking Action: Epistemologies, Methodologies, and Methods in Qualitative Research. *Qualitative Health Research*, *17*(10), 1316–1328. https://doi.org/10.1177/1049732307306927

Casadevall, A., Steen, R. G., & Fang, F. C. (2014). Sources of error in the retracted scientific literature. *The FASEB Journal*, *28*(9), 3847–3855. https://doi.org/10.1096/fj.14-256735

Chambers, C. D. (2013). Registered Reports: A new publishing initiative at Cortex. *Cortex*, *49*(3), 609–610. https://doi.org/10.1016/j.cortex.2012.12.016

Chan, A.-W., Hróbjartsson, A., Haahr, M. T., Gøtzsche, P. C., & Altman, D. G. (2004). Empirical Evidence for Selective Reporting of Outcomes in Randomized Trials: Comparison of Protocols to Published Articles. *JAMA*, *291*(20), 2457. https://doi.org/10.1001/jama.291.20.2457

Chen, X. (1994). The rule of reproducibility and its applications in experiment appraisal. *Synthese*, *99*(1), 87–109. https://doi.org/10.1007/BF01064532

Chmielewski, M., & Kucker, S. C. (2020). An MTurk Crisis? Shifts in Data Quality and the Impact on Study Results. *Social Psychological and Personality Science*, *11*(4), 464–473. https://doi.org/10.1177/1948550619875149

Choi, H. H. (2024). Preregistration and Predictivism. *Synthese*, *204*(6), 1–22.

Chubin, D. E. (1985). Open Science and Closed Science: Tradeoffs in a Democracy. *Science, Technology, & Human Values*, *10*(2), 73–80. https://doi.org/10.1177/016224398501000211

Ciubotariu, I. I., & Bosch, G. (2022). Improving research integrity: A framework for responsible science communication. *BMC Research Notes*, *15*(1), 177. https://doi.org/10.1186/s13104-022-06065-5

Claerbout, J. F., & Karrenbach, M. (1992). Electronic documents give reproducible research a new meaning. *SEG Technical Program Expanded Abstracts 1992*, 601–604. https://doi.org/10.1190/1.1822162

Claesen, A., Gomes, S., Tuerlinckx, F., & Vanpaemel, W. (2021). Comparing dream to reality: an assessment of adherence of the first generation of preregistered studies. Royal Society Open Science, 8(10). https://doi.org/10.1098/rsos.211037

Clancy, R. F., Lee, L. M., Zhu, Q., Plemmons, D., Heitman, E., McIntosh, T., Kalichman, M., Thrush, C., Grossenbacher, L., Williams, B., Zhu, M., &

Jenkins, I. (2026). Toward an “ecosystems” approach to responsible conduct of research (RCR): A multi-stakeholder framework for collaborative accountability and policy recommendations on research integrity. *Accountability in Research*, *33*(4), 2596906. https://doi.org/10.1080/08989621.2025.2596906

Clarke, B., Lee, P. Y. (Katherine), Schiavone, S. R., Rhemtulla, M., & Vazire, S. (2024). The prevalence of direct replication articles in top-ranking psychology journals. *American Psychologist*. https://doi.org/10.1037/amp0001385

COARRA. (2025). *Transitioning to Open Infrastructures for Responsible Research Assessment: Barriers, Enablers, and Strategic Recommendations*. Zenodo. https://doi.org/10.5281/ZENODO.17399516

Cole, N. L., Ulpts, S., Bochynska, A., Kormann, E., Good, M., Leitner, B., & Ross-Hellauer, T. (2024). *Reproducibility and replicability of qualitative research: An integrative review of concepts, barriers and enablers*. https://doi.org/10.31222/osf.io/n5zkw_v1

Collins, H. (2016). Reproducibility of Experiments: Experimenters’ Regress, Statistical Uncertainty Principle, and the Replication Imperative. In H. Atmanspacher & S. Maasen (Eds.), *Reproducibility: Principles, Problems, Practices, and Prospects* (pp. 65–82). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118865064.ch4

Collins, H. (2021). Science as craftwork with integrity. In D. Ludwig, I. Koskinen, Z. Mncube, L. Poliseli, & L. Reyes-Galindo, *Global Epistemologies and Philosophies of Science* (1st ed., pp. 296–307). Routledge. https://doi.org/10.4324/9781003027140-32

Collins, H., & Evans, R. (2007). *Rethinking Expertise*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226113623.001.0001

Collins, H. M. (1974). The TEA Set: Tacit Knowledge and Scientific Networks. *Science Studies*, *4*(2), 165–185. https://doi.org/10.1177/030631277400400203

Collins, H. M. (1975). The Seven Sexes: A Study in the Sociology of a Phenomenon, or the Replication of Experiments in Physics. *Sociology*, *9*(2), 205–224. https://doi.org/10.1177/003803857500900202

Collins, H. M. (1976, November 4-6). *Upon the replication of scientific findings: A discussion illuminated by the experiences of researchers into parapsychology* [Paper presentation]. Society for Social Studies of Science 4S/ISA 1976 Conference, Ithaca, NY, USA

Collins, H. M. (1981). Introduction: Stages in the Empirical Programme of Relativism. *Social Studies of Science*, *11*(1), 3–10. JSTOR.

Collins, H. M. (1985). *Changing order: Replication and induction in scientific practice*. Sage Beverly Hills, Calif.

Collins, H. M. (1991). The Meaning of Replication and the Science of Economics. *History of Political Economy*, *23*(1), 123–142. https://doi.org/10.1215/00182702-23-1-123

Collins, H. M., & Evans, R. (2002). The Third Wave of Science Studies: Studies of Expertise and Experience. *Social Studies of Science*, *32*(2), 235–296. https://doi.org/10.1177/0306312702032002003
Collins, H., & Mills, M. A. (2022). Science as Craftwork with Integrity. *The New Atlantis*, (67), 53–70. JSTOR.
Collins, H., Shrager, J., Bartlett, A., Conley, S., Hale, R., & Evans, R. (2023). Hypernormal Science and its Significance. *Perspectives on Science*, *31*(2), 262–292. https://doi.org/10.1162/posc_a_00572
Conry-Murray, C., Mcconnon, A., & Bower, M. (2022). The Effect of Preregistration and P-Value Patterns on Trust in Psychology and Biology Research. *Collabra: Psychology*, *8*(1), 36306. https://doi.org/10.1525/collabra.36306
Corker, K. S., Waltman, L., & Coates, J. A. (2024). *Understanding the Publish-Review-Curate (PRC) Model of Scholarly Communication*. MetaArXiv. https://doi.org/10.31222/osf.io/h7swt
Cova, F., Strickland, B., Abatista, A., Allard, A., Andow, J., Attie, M., Beebe, J., Berniūnas, R., Boudesseul, J., Colombo, M., Cushman, F., Diaz, R., N'Djaye Nikolai Van Dongen, N., Dranseika, V., Earp, B. D., Torres, A. G., Hannikainen, I., Hernández-Conde, J. V., Hu, W., … Zhou, X. (2021). Correction to: Estimating the Reproducibility of Experimental Philosophy. *Review of Philosophy and Psychology*, *12*(1), 45–48. https://doi.org/10.1007/s13164-018-0407-2
Cramer, M., & Reinhart, M. (2026). Governing Science through Evaluation: A Global Heuristic of Research Evaluation Regimes. *Minerva*. https://doi.org/10.1007/s11024-026-09633-2
Crane, D. (1972). *Invisible colleges; diffusion of knowledge in scientific communities*. University of Chicago Press.
Cyranoski, D. (2014). Stem-cell scientist found guilty of misconduct. *Nature*, nature.2014.14974. https://doi.org/10.1038/nature.2014.14974
Dal-Ré, R., Bouter, L. M., Cuijpers, P., Gluud, C., & Holm, S. (2020). Should research misconduct be criminalized? *Research Ethics*, *16*(1–2), 1–12. https://doi.org/10.1177/1747016119898400
Dal-Ré, R., & Marušić, A. (2025). The definition of research misconduct should be stated in the abstract when reporting research on research misconduct. *Accountability in Research*, *32*(4), 639–647. https://doi.org/10.1080/08989621.2024.2306538
Danziger, K. (1985). The Methodological Imperative in Psychology. *Philosophy of the Social Sciences*, *15*(1), 1–13. https://doi.org/10.1177/004839318501500101
Daston, L. (1995). The Moral Economy of Science. *Osiris*, *10*, 2–24. https://doi.org/10.1086/368740
Daston, L., & Galison, P. (1992). The Image of Objectivity. *Representations*, *40*, 81–128. https://doi.org/10.2307/2928741
Daston, L., & Galison, P. (2007). *Objectivity*. Zone Books. JSTOR. https://doi.org/10.2307/j.ctv1c9hq4d

Davies, S. R. (2019). An Ethics of the System: Talking to Scientists About Research Integrity. *Science and Engineering Ethics*, *25*(4), 1235–1253. https://doi.org/10.1007/s11948-018-0064-y

Davies, S. R., & Lindvig, K. (2021). Assembling research integrity: Negotiating a policy object in scientific governance. *Critical Policy Studies*, *15*(4), 444–461. https://doi.org/10.1080/19460171.2021.1879660

Davin, C., Beck, J., & Ma, L. (2026). Open Practices, Closed Realities? Archaeological Perspectives on Open Research Practices. *The Journal of Electronic Publishing*, *29*(1). https://doi.org/10.3998/jep.7833

De Angelis, C., Drazen, J. M., Frizelle, F. A., Haug, C., Hoey, J., Horton, R., Kotzin, S., Laine, C., Marusic, A., Overbeke, A. J. P., Schroeder, T. V., Sox, H. C., & Van Der Weyden, M. B. (2004). Clinical trial registration: A statement from the International Committee of Medical Journal Editors. *The Lancet*, *364*(9438), 911–912. https://doi.org/10.1016/S0140-6736(04)17034-7

De Peuter, S., & Conix, S. (2023). Fostering a research integrity culture: Actionable advice for institutions. *Science and Public Policy*, *50*(1), 133–145. https://doi.org/10.1093/scipol/scac059

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. https://doi.org/10.1525/jer.2006.1.1.43

Degn, L. (2020). Integrating Integrity: The Organizational Translation of Policies on Research Integrity. *Science and Engineering Ethics*, *26*(6), 3167–3182. https://doi.org/10.1007/s11948-020-00262-w

DeHaven, A. (2017, May 23). *Preregistration: A Plan, Not a Prison*. Retrieved September 1, 2025, from https://www.cos.io/blog/preregistration-plan-not-prison

Delamont, S., & Atkinson, P. (2001). Doctoring Uncertainty: Mastering Craft Knowledge. *Social Studies of Science*, *31*(1), 87–107. https://doi.org/10.1177/030631201031001005

Derksen, M. (2019). Putting Popper to work. *Theory & Psychology*, *29*(4), 449–465. https://doi.org/10.1177/0959354319838343

Derksen, M. (2021). A Menagerie of Imposters and Truth-Tellers: Diederik Stapel and the Crisis in Psychology. In M. Derksen, *The Imposter as Social Theory* (pp. 53–76). Policy Press. https://doi.org/10.1332/policypress/9781529213072.003.0003

Derksen, M., & Field, S. (2022). The Tone Debate: Knowledge, Self, and Social Order. *Review of General Psychology*, *26*(2), 172–183. https://doi.org/10.1177/10892680211015636

Derksen, M., Meirmans, S., Brenninkmeijer, J., Pols, J., De Boer, A., Van Eyghen, H., Gayet, S., Groenwold, R., Hernaus, D., Huijnen, P., Jonker, N., De Kleijn, R., Kroll, C. F., Krypotos, A.-M., Van Der Laan, N., Luijken, K., Meijer, E., Pear, R. S. A., Peels, R., … De Winter, J. (2025). Replication studies in the Netherlands: Lessons learned and recommendations for funders, publishers

and editors, and universities. *Accountability in Research*, *32*(7), 1285–1303. https://doi.org/10.1080/08989621.2024.2383349
Derksen, M., & Morawski, J. (2022). Kinds of Replication: Examining the Meanings of "Conceptual Replication" and "Direct Replication." *Perspectives on Psychological Science*, *17*(5), 1490–1505. https://doi.org/10.1177/17456916211041116
Derksen, M., & Morawski, J. (2024). Replications are informative, particularly when they fail. *Theory & Psychology*, *34*(5), 597–603. https://doi.org/10.1177/09593543241269697
Desmond, H., & Dierickx, K. (2021). Research integrity codes of conduct in Europe: Understanding the divergences. *Bioethics*, *35*(5), 414–428. https://doi.org/10.1111/bioe.12851
Deutsche Forschungsgemeinschaft. (2019). *Guidelines for Safeguarding Good Research Practice. Code of Conduct*. https://doi.org/10.5281/ZENODO.3923602
Deutsche Forschungsgemeinschaft. (2025). *Guidelines for Safeguarding Good Research Practice. Code of Conduct*. https://doi.org/10.5281/ZENODO.14281892
Deutsche Forschungsgemeinschaft. (2026). *Sicherung guter wissenschaftlicher Praxis. Denkschrift / Safeguarding Good Scientific Practice. Memorandum (1998)*. Deutsche Forschungsgemeinschaft (DFG). https://doi.org/10.5281/ZENODO.17986720
Devezer, B., Navarro, D. J., Vandekerckhove, J., & Ozge Buzbas, E. (2021). The case for formal methodology in scientific reform. *Royal Society Open Science*, *8*(3), rsos.200805, 200805. https://doi.org/10.1098/rsos.200805
Devezer, B., & Penders, B. (2023). Scientific reform, citation politics and the bureaucracy of oblivion. *Quantitative Science Studies*, *4*(4), 857–859. https://doi.org/10.1162/qss_c_00274
Dhai, A. (2014). The research ethics evolution: From Nuremberg to Helsinki. *South African Medical Journal*, *104*(3), 178. https://doi.org/10.7196/SAMJ.7864
Diaba-Nuhoho, P., & Amponsah-Offeh, M. (2021). Reproducibility and research integrity: The role of scientists and institutions. *BMC Research Notes*, *14*(1), 451. https://doi.org/10.1186/s13104-021-05875-3
Dirnagl, U. (2020). Preregistration of exploratory research: Learning from the golden age of discovery. *PLOS Biology*, *18*(3), e3000690. https://doi.org/10.1371/journal.pbio.3000690
Donoho, D. L., Maleki, A., Rahman, I. U., Shahram, M., & Stodden, V. (2009). Reproducible Research in Computational Harmonic Analysis. *Computing in Science & Engineering*, *11*(1), 8–18. https://doi.org/10.1109/MCSE.2009.15
Douglas, H. (2024). Structuring Institutions for Responsible and Accountable Science. *Philosophy of Science*, *91*(5), 1149–1158. https://doi.org/10.1017/psa.2023.128
Douglas, H., & Branch, T. Y. (2024). The social contract for science and the value-free ideal. *Synthese*, *203*(2), 40. https://doi.org/10.1007/s11229-023-04477-9

Douglas, M. (1975). *Implicit meanings: Essays in anthropology*. Routledge and Kegan Paul. WorldCat.

Douglas-Jones, R., & Wright, S. (2017). *Mapping the Integrity Landscape: Organisations, Policies, Concepts* (pp. 1–35) [WorkingPaper]. CHEF, Danish School of Education, Aarhus University.

Downey, F., & Veitch, R. (2021). Research Integrity—It's about More than Misconduct. *Forensic Sciences Research*, *6*(4), 283–284. https://doi.org/10.1080/20961790.2021.1971383

Duncombe, J., & Jessop, J. (2002). 'Doing Rapport' and the Ethics of 'Faking Friendship.' In M. Mauthner, M. Birch, J. Jessop, & T. Miller, *Ethics in Qualitative Research* (pp. 108–122). SAGE Publications Ltd. https://doi.org/10.4135/9781849209090.n6

Durkheim, É. (1982). *The rules of sociological method* (1. ed). Free Press.

Earp, B. D., & Trafimow, D. (2015). Replication, falsification, and the crisis of confidence in social psychology. *Frontiers in Psychology*, *6*. https://doi.org/10.3389/fpsyg.2015.00621

Ebersole, C. R., Mathur, M. B., Baranski, E., Bart-Plange, D.-J., Buttrick, N. R., Chartier, C. R., Corker, K. S., Corley, M., Hartshorne, J. K., IJzerman, H., Lazarević, L. B., Rabagliati, H., Ropovik, I., Aczel, B., Aeschbach, L. F., Andrighetto, L., Arnal, J. D., Arrow, H., Babincak, P., … Nosek, B. A. (2020). Many Labs 5: Testing Pre-Data-Collection Peer Review as an Intervention to Increase Replicability. *Advances in Methods and Practices in Psychological Science*, *3*(3), 309–331. https://doi.org/10.1177/2515245920958687

Edlund, J. E., Cuccolo, K., Irgens, M. S., Wagge, J. R., & Zlokovich, M. S. (2022). Saving Science Through Replication Studies. *Perspectives on Psychological Science*, *17*(1), 216–225. https://doi.org/10.1177/1745691620984385

Edwards, M. A., & Roy, S. (2017). Academic Research in the 21st Century: Maintaining Scientific Integrity in a Climate of Perverse Incentives and Hypercompetition. *Environmental Engineering Science*, *34*(1), 51–61. https://doi.org/10.1089/ees.2016.0223

Elizondo, A., & Kaltenbrunner, W. (2024). Navigating the Science System: Research Integrity and Academic Survival Strategies. *Science and Engineering Ethics*, *30*(2), 12. https://doi.org/10.1007/s11948-024-00467-3

Elliott, K. C. (2022). A Taxonomy of Transparency in Science. *Canadian Journal of Philosophy*, *52*(3), 342–355. https://doi.org/10.1017/can.2020.21

Emic/etic distinction. (2008). In L. Given, *The SAGE Encyclopedia of Qualitative Research Methods*. SAGE Publications, Inc. https://doi.org/10.4135/9781412963909.n130

Emons, W., Sijtsma, K., & Bouter, L. (2025). *Preregistration of Research on Research Integrity is still Not Common: Findings from the Hong Kong, Cape Town, and Athens Editions of the World Conference on Research Integrity*. https://doi.org/10.31222/osf.io/cn6jf

Errington, T. M., Denis, A., Perfito, N., Iorns, E., & Nosek, B. A. (2021). Challenges for assessing replicability in preclinical cancer biology. *eLife*, *10*, e67995. https://doi.org/10.7554/eLife.67995

Errington, T. M., Mathur, M., Soderberg, C. K., Denis, A., Perfito, N., Iorns, E., & Nosek, B. A. (2021). Investigating the replicability of preclinical cancer biology. *eLife*, *10*, e71601. https://doi.org/10.7554/eLife.71601

Fanelli, D. (2011). The black, the white and the grey areas—Towards an international and interdisciplinary definition of scientific misconduct. In N. Steneck & T. Meyer (Eds.), *Promoting Research Integrity on a Global Basis*.

Fanelli, D. (2018). Is science really facing a reproducibility crisis, and do we need it to? *Proceedings of the National Academy of Sciences*, *115*(11), 2628–2631. https://doi.org/10.1073/pnas.1708272114

Fecher, B., & Friesike, S. (2014). Open Science: One Term, Five Schools of Thought. In S. Bartling & S. Friesike (Eds.), *Opening Science* (pp. 17–47). Springer International Publishing. https://doi.org/10.1007/978-3-319-00026-8_2

Feenstra, R. A., Carretero García, C., & Gómez Nicolau, E. (2025). Perception of Research Misconduct in a Spanish University. *Journal of Academic Ethics*, *23*(1), 89–112. https://doi.org/10.1007/s10805-024-09526-7

Feest, U. (2016). The experimenters' regress reconsidered: Replication, tacit knowledge, and the dynamics of knowledge generation. *Studies in History and Philosophy of Science Part A*, *58*, 34–45. https://doi.org/10.1016/j.shpsa.2016.04.003

Feest, U. (2019). Why Replication Is Overrated. *Philosophy of Science*, *86*(5), 895–905. https://doi.org/10.1086/705451

Felt, U. (2017). "Response-able Practices" or "New Bureaucracies of Virtue": The Challenges of Making RRI Work in Academic Environments. In L. Asveld, R. Van Dam-Mieras, T. Swierstra, S. Lavrijssen, K. Linse, & J. Van Den Hoven (Eds.), *Responsible Innovation 3* (pp. 49–68). Springer International Publishing. https://doi.org/10.1007/978-3-319-64834-7_4

Felt, U., & Frantz, F. (2026). The Shifting 'Self' of Science's Self-Governing Capacity: Four Decades of Research Integrity Discussions in *Science* and *Nature*. *Social Studies of Science*, *56*(2), 206–232. https://doi.org/10.1177/03063127251392603

Ferguson, C. J., & Heene, M. (2012). A Vast Graveyard of Undead Theories: Publication Bias and Psychological Science's Aversion to the Null. *Perspectives on Psychological Science*, *7*(6), 555–561. https://doi.org/10.1177/1745691612459059

Fernández Pinto, M. (2020). Open Science for private Interests? How the Logic of Open Science Contributes to the Commercialization of Research. *Frontiers in Research Metrics and Analytics*, *5*, 588331. https://doi.org/10.3389/frma.2020.588331

Festinger, L. (1957). *A theory of cognitive dissonance*. Stanford University Press.

Feyerabend, P. (1976). *Against method: Outline of an anarchistic theory of knowledge*.
Humanities Press.
Feynman, R. P. (1974). Cargo cult science: Some remarks on science, pseudoscience, and learning how to not fool yourself. *Engineering and Science*, *37*(7), 10–13.
Field, S. M. (2022). *Charting the Constellation of Science Reform*. https://doi.org/10.31219/osf.io/udfw4
Field, S. M., & Derksen, M. (2021). Experimenter as automaton; experimenter as human: Exploring the position of the researcher in scientific research. *European Journal for Philosophy of Science*, *11*(1), 11. https://doi.org/10.1007/s13194-020-00324-7
Field, S. M., Kiers, H., & Derksen, M. (2024). *Exploring the Constellation of Communities Shaping Science Reform and its Future Progress*. Open Science Framework. https://doi.org/10.31219/osf.io/5up7c
Field, S. M., Thompson, J., De Rijcke, S., Penders, B., & Munafò, M. R. (2024). Exploring the dimensions of responsible research systems and cultures: A scoping review. *Royal Society Open Science*, *11*(1), 230624. https://doi.org/10.1098/rsos.230624
Field, S. M., Wagenmakers, E.-J., Kiers, H. A. L., Hoekstra, R., Ernst, A. F., & Van Ravenzwaaij, D. (2020). The effect of preregistration on trust in empirical research findings: Results of a registered report. *Royal Society Open Science*, *7*(4), 181351. https://doi.org/10.1098/rsos.181351
Fleck, L. (1935). *Entstehung und Entwicklung einer wissenschaftlichen Tatsache. Einführung in die Lehre vom Denkstil und Denkkollektiv, etc*. WorldCat.
Fletcher, S. C. (2021). The role of replication in psychological science. *European Journal for Philosophy of Science*, *11*(1), 1–19.
Flis, I. (2019). Psychologists psychologizing scientific psychology: An epistemological reading of the replication crisis. *Theory & Psychology*, *29*(2), 158–181. https://doi.org/10.1177/0959354319835322
Flis, I. (2022). The Function of Literature in Psychological Science. *Review of General Psychology*, *26*(2), 146–156. https://doi.org/10.1177/10892680211066466
Folk, D., & Dunn, E. (2023). A systematic review of the strength of evidence for the most commonly recommended happiness strategies in mainstream media. *Nature Human Behaviour*, *7*(10), 1697–1707. https://doi.org/10.1038/s41562-023-01651-4
Forsberg, E.-M., Anthun, F. O., Bailey, S., Birchley, G., Bout, H., Casonato, C., Fuster, G. G., Heinrichs, B., Horbach, S., Jacobsen, I. S., Janssen, J., Kaiser, M., Lerouge, I., Van Der Meulen, B., De Rijcke, S., Saretzki, T., Sutrop, M., Tazewell, M., Varantola, K., … Zöller, M. (2018). Working with Research Integrity—Guidance for Research Performing Organisations: The Bonn PRINTEGER Statement. *Science and Engineering Ethics*, *24*(4), 1023–1034. https://doi.org/10.1007/s11948-018-0034-4

Francis, J. R. (1989). The credibility and legitimation of science: A loss of faith in the scientific narrative. *Accountability in Research*, *1*(1), 5–22. https://doi.org/10.1080/08989628908573770

Frank, M., Miller, B., Vosskuhl, J., Zänkert, S., Heßelmann, F., & Jolliffe, J. (2026). Error needs culture! *Exploring the relationship between error culture and openness*. *Royal Society Open Science*, *13*(2), 242233. https://doi.org/10.1098/rsos.242233

Franklin, A., & Collins, H. (2016). Two Kinds of Case Study and a New Agreement. In R. Scholl & T. Sauer (Eds.), *The Philosophy of Historical Case Studies* (pp. 95–121). Springer Verlag.

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing* (1st ed.). Oxford University PressOxford. https://doi.org/10.1093/acprof:oso/9780198237907.001.0001

Fujimura, J. H. (1992). Crafting science: Standardized packages, boundary objects, and "translation." In A. Pickering (Ed.), *Science as practice and culture* (pp. 168--211). University of Chicago Press.

Galison, P., & Stump, D. J. (Eds.). (1996). *The Disunity of science: Boundaries, contexts, and power*. Stanford University Press.

Garcia-Sancho, M., & Knorr Cetina, K. (2018). "These Were Not Boring Meetings": Miguel García-Sancho Talks with Karin Knorr Cetina. *Engaging Science, Technology, and Society*, *4*, 246–266. https://doi.org/10.17351/ests2018.239

Garisto, D. (2023). LK-99 isn't a superconductor—How science sleuths solved the mystery. *Nature*, *620*(7975), 705–706. https://doi.org/10.1038/d41586-023-02585-7

Gärtner, A., Leising, D., & Schönbrodt, F. D. (2024). Towards responsible research assessment: How to reward research quality. *PLOS Biology*, *22*(2), e3002553. https://doi.org/10.1371/journal.pbio.3002553

Garud, R., Gehman, J., & Karunakaran, A. (2014). Boundaries, breaches, and bridges: The case of Climategate. *Research Policy*, *43*(1), 60–73. https://doi.org/10.1016/j.respol.2013.07.007

Geertz, C. (1973). Thick Description: Toward an Interpretive Theory of Culture. In *The Interpretation of Cultures: Selected Essays* (pp. 3–30). New York: Basic Books.

Geiger, B. B. (2021). Performing Trustworthiness: The 'Credibility Work' of Prominent Sociologists. *Sociology*, *55*(4), 785–802. https://doi.org/10.1177/0038038520977805

Gelman, A., & Loken, E. (2014). The Statistical Crisis in Science. *American Scientist*, *102*(6), 460. https://doi.org/10.1511/2014.111.460

Gergen, K. J. (1973). Social psychology as history. *Journal of Personality and Social Psychology*, *26*(2), 309–320. https://doi.org/10.1037/h0034436

Gieryn, T. F. (1983). Boundary-Work and the Demarcation of Science from Non-Science: Strains and Interests in Professional Ideologies of Scientists. *American Sociological Review*, *48*(6), 781. https://doi.org/10.2307/2095325

Gieryn, T. F. (1995). Boundaries of Science. In S. Jasanoff, G. Markle, J. Peterson, & T. Pinch, *Handbook of Science and Technology Studies* (pp. 392–443). SAGE Publications, Inc. https://doi.org/10.4135/9781412990127.n18
Gieryn, T. F. (1999). *Cultural boundaries of science: Credibility on the line*. University of Chicago Press.
Gigerenzer, G. (2004). Mindless statistics. *The Journal of Socio-Economics*, *33*(5), 587–606. https://doi.org/10.1016/j.socec.2004.09.033
Gilbert, G. N. (1977). Referencing as Persuasion. *Social Studies of Science*, *7*(1), 113–122. http://www.jstor.org/stable/284636
Gilbert, G. N., & Mulkay, M. (1984). *Opening Pandora's Box: A Sociological Analysis of Scientists' Discourse* (Issue 1). CUP Archive.
Gläser, J. (2024). How can we make 'research quality' a theoretical concept? *Research Evaluation*, *33*, rvae038. https://doi.org/10.1093/reseval/rvae038
Gläser, J., Bielick, J., Jungmann, R., Laudel, G., Lettkemann, E., Petschick, G., & Tschida, U. (2015). Research cultures as an explanatory factor. *Österreichische Zeitschrift Für Soziologie*, *40*(3), 327–346. https://doi.org/10.1007/s11614-015-0177-3
Goffman, E. (1983). The Interaction Order: American Sociological Association, 1982 Presidential Address. *American Sociological Review*, *48*(1), 1–17. https://doi.org/10.2307/2095141
Goffman, E. (1989). ON FIELDWORK. *Journal of Contemporary Ethnography*, *18*(2), 123–132. https://doi.org/10.1177/089124189018002001
Gopalakrishna, G., Ter Riet, G., Vink, G., Stoop, I., Wicherts, J. M., & Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in The Netherlands. *PLOS ONE*, *17*(2), e0263023. https://doi.org/10.1371/journal.pone.0263023
Grant, S., Corker, K. S., Mellor, D. T., Stewart, S. L. K., Cashin, A. G., Lagisz, M., Mayo-Wilson, E., Moher, D., Umpierre, D., Barbour, V., Buck, S., Collins, G., Hazlett, H., Hrynaszkiewicz, I., Lee, C. J., Parker, T. H., Rethlefsen, M. L., Toomey, E., & Nosek, B. A. (2025). *TOP 2025: An Update to the Transparency and Openness Promotion Guidelines*. https://doi.org/10.31222/osf.io/nmfs6_v2
Green, C. D. (2015). Why Psychology Isn't Unified, and Probably Never Will Be. *Review of General Psychology*, *19*(3), 207–214. https://doi.org/10.1037/gpr0000051
Guillemin, M., & Gillam, L. (2004). Ethics, Reflexivity, and "Ethically Important Moments" in Research. *Qualitative Inquiry*, *10*(2), 261–280. https://doi.org/10.1177/1077800403262360
Gulbrandsen, J. M. (2000). *Research quality and organisational factors: An investigation of the relationship*. Trondheim: NTNU.
Gundersen, O. E. (2021). The fundamental principles of reproducibility. *Philosophical Transactions of the Royal Society A: Mathematical, Physical*

and Engineering Sciences, *379*(2197), 20200210. https://doi.org/10.1098/rsta.2020.0210
Guston, D. H. (1994). The demise of the social contract for science: Misconduct in science and the nonmodern world. *The Centennial Review*, *38*(2), 215–248.
Guston, D. H. (2000). *Between politics and science: Assuring the integrity and productivity of research*. Cambridge University Press.
Guttinger, S. (2018). Replications Everywhere: Why the replication crisis might be less severe than it seems at first. *BioEssays*, *40*(7), 1800055. https://doi.org/10.1002/bies.201800055
Guttinger, S. (2019). A New Account of Replication in the Experimental Life Sciences. *Philosophy of Science*, *86*(3), 453–471. https://doi.org/10.1086/703555
Guttinger, S. (2020). The limits of replicability. *European Journal for Philosophy of Science*, *10*(2), 10. https://doi.org/10.1007/s13194-019-0269-1
Hacking, I. (2007). Kinds of People: Moving Targets. In H. Ian (Ed.), *Proceedings of the British Academy, Volume 151, 2006 Lectures* (Vol. 151, pp. 285–318).
Hagmann, M. (2000). Panel Finds Scores of Suspect Papers in German Fraud Probe. *Science*, *288*(5474), 2106–2107. https://doi.org/10.1126/science.288.5474.2106
Hagstrom, W. O. (1965). *The scientific community*. Basic Books
Haig, B. D. (2022). Understanding Replication in a Way That Is True to Science. *Review of General Psychology*, *26*(2), 224–240. https://doi.org/10.1177/10892680211046514
Halffman, W., & Radder, H. (2015). The Academic Manifesto: From an Occupied to a Public University. *Minerva*, *53*(2), 165–187. https://doi.org/10.1007/s11024-015-9270-9
Hammersley, M., & Atkinson, P. (2019). *Ethnography: Principles in Practice* (4th ed.). Routledge. https://doi.org/10.4324/9781315146027
Hangel, N., & Schickore, J. (2017). Scientists' Conceptions of Good Research Practice. *Perspectives on Science*, *25*(6), 766–791. https://doi.org/10.1162/POSC_a_00265
Hanson, M. A., Barreiro, P. G., Crosetto, P., & Brockington, D. (2024). The strain on scientific publishing. *Quantitative Science Studies*, *5*(4), 823–843. https://doi.org/10.1162/qss_a_00327
Haraway, D. (1988). Situated Knowledges: The Science Question in Feminism and the Privilege of Partial Perspective. *Feminist Studies*, *14*(3), 575. https://doi.org/10.2307/3178066
Hardwicke, T. E., Serghiou, S., Janiaud, P., Danchev, V., Crüwell, S., Goodman, S. N., & Ioannidis, J. P. A. (2020). Calibrating the Scientific Ecosystem Through Meta-Research. Annual Review of Statistics and Its Application, 7(1), 11–37. https://doi.org/10.1146/annurev-statistics-031219-041104
Hardwicke, T. E., & Wagenmakers, E.-J. (2023). Reducing bias, increasing transparency and calibrating confidence with preregistration. *Nature Human Behaviour*, *7*(1), 15–26. https://doi.org/10.1038/s41562-022-01497-2

Hastings, R., Labib, K., Lechner, I., Bouter, L., Widdershoven, G., & Evans, N. (2023). Guidance on research integrity provided by pan-European discipline-specific learned societies: A scoping review. *Science and Public Policy*, *50*(2), 318–335. https://doi.org/10.1093/scipol/scac067

Haven, T. (2025). It takes two flints to start a fire: A focus group study into PhD supervision for responsible research. *Accountability in Research*, *32*(5), 717–740. https://doi.org/10.1080/08989621.2025.2457584

Haven, T., Bouter, L., Mennen, L., & Tijdink, J. (2023). Superb supervision: A pilot study on training supervisors to convey responsible research practices onto their PhD candidates. *Accountability in Research*, *30*(8), 574–591. https://doi.org/10.1080/08989621.2022.2071153

Haven, T., Gopalakrishna, G., Tijdink, J., Van Der Schot, D., & Bouter, L. (2022). Promoting trust in research and researchers: How open science and research integrity are intertwined. *BMC Research Notes*, *15*(1), 302. https://doi.org/10.1186/s13104-022-06169-y

Haven, T., Pasman, H. R., Widdershoven, G., Bouter, L., & Tijdink, J. (2020). Researchers' Perceptions of a Responsible Research Climate: A Multi Focus Group Study. *Science and Engineering Ethics*, *26*(6), 3017–3036. https://doi.org/10.1007/s11948-020-00256-8

Helgesson, G., & Bülow, W. (2023). Research Integrity and Hidden Value Conflicts. *Journal of Academic Ethics*, *21*(1), 113–123. https://doi.org/10.1007/s10805-021-09442-0

Hendriks, B., & Reinhart, M. (2020). Science Blogs as Critique—Building Public Identities in the Field of Translational Research. *Science & Technology Studies*, *33*(3), 19–38. https://doi.org/10.23987/sts.75153

Herzfeld, M. (2023). Ethnographic responsibility: Can the bureaucratization of research ethics be ethical? *Anthropology Today*, *39*(3), 3–6. https://doi.org/10.1111/1467-8322.12811

Hess, D. (2001). Ethnography and the Development of Science and Technology Studies. In P. Atkinson, A. Coffey, S. Delamont, J. Lofland, & L. Lofland, *Handbook of Ethnography* (pp. 234–245). SAGE Publications Ltd. https://doi.org/10.4135/9781848608337.n16

Hesselmann, F., Graf, V., Schmidt, M., & Reinhart, M. (2017). The visibility of scientific misconduct: A review of the literature on retracted journal articles. *Current Sociology*, *65*(6), 814–845. https://doi.org/10.1177/0011392116663807

Hesselmann, F., & Reinhart, M. (2021). Cycles of invisibility: The limits of transparency in dealing with scientific misconduct. *Social Studies of Science*, *51*(3), 414–438. https://doi.org/10.1177/0306312720975201

Hesselmann, F., & Reinhart, M. (2024). From scandal to reform: Approaches to research integrity at a turning point. *Journal of Responsible Innovation*, *11*(1), 2414491. https://doi.org/10.1080/23299460.2024.2414491

Hesselmann, F., Wienefoet, V., & Reinhart, M. (2014). Measuring Scientific Misconduct—Lessons from Criminology. *Publications*, *2*(3), 61–70. https://doi.org/10.3390/publications2030061

Hessels, L. K., Franssen, T., Scholten, W., & De Rijcke, S. (2019). Variation in Valuation: How Research Groups Accumulate Credibility in Four Epistemic Cultures. *Minerva*, *57*(2), 127–149. https://doi.org/10.1007/s11024-018-09366-x

Hessels, L. K., Van Lente, H., & Smits, R. (2009). In search of relevance: The changing contract between science and society. *Science and Public Policy*, *36*(5), 387–401. https://doi.org/10.3152/030234209X442034

Hochbruck, M. (2024). Überarbeitung der DFG-Denkschrift »Sicherung guter wissenschaftlicher Praxis«. *Wissenschaftsrecht*, *57*(1), 13–21. https://doi.org/10.1628/wissr-2024-0004

Hoekstra, R., & Vazire, S. (2021). Aspiring to greater intellectual humility in science. *Nature Human Behaviour*, *5*(12), 1602–1607. https://doi.org/10.1038/s41562-021-01203-8

Horbach, S. P. J. M., Cole, N. L., Kopeinik, S., Leitner, B., Ross-Hellauer, T., & Tijdink, J. (2026). How to Get There From Here? Barriers and Enablers on the Road Towards Reproducibility in Research. *Journal of Trial and Error*. https://doi.org/10.36850/f4c4-4a38

Horbach, S. P. J. M., Fishberg, R., Ulpts, S., & Degn, L. (2024). Thou Shalt Not! – How the institutional afterlife of research misconduct scandals shapes research integrity training. *Journal of Responsible Innovation*, *11*(1), 2414500. https://doi.org/10.1080/23299460.2024.2414500

Horbach, S. P. J. M., & Halffman, W. (2017). Promoting Virtue or Punishing Fraud: Mapping Contrasts in the Language of 'Scientific Integrity.' *Science and Engineering Ethics*, *23*(6), 1461–1485. https://doi.org/10.1007/s11948-016-9858-y

Horbach, S. P. M. J., & Valinciute, A. (2026). Contextualizing the emerging and contested role of sleuths in the science monitoring system. *Science and Public Policy*, scag044. https://doi.org/10.1093/scipol/scag044

Hosseini, M., Senabre Hidalgo, E., Horbach, S. P. J. M., Güttinger, S., & Penders, B. (2024). Messing with Merton: The intersection between open science practices and Mertonian values. *Accountability in Research*, *31*(5), 428–455. https://doi.org/10.1080/08989621.2022.2141625

Hostler, T. (2024a). Research assessment using a narrow definition of "research quality" is an act of gatekeeping: A comment on Gärtner et al. (2022). *Meta-Psychology*, *8*. https://doi.org/10.15626/MP.2023.3764

Hostler, T. J. (2024b). Open research reforms and the capitalist university: Areas of opposition and alignment. *Collabra: Psychology*, *10*(1), 121383. https://doi.org/10.1525/collabra.121383

Hostler, T. J. (2023). The Invisible Workload of Open Research. *Journal of Trial and Error*, *4*(1), 21–36. https://doi.org/10.36850/mr5

Hudson, R. (2023). Explicating Exact versus Conceptual Replication. *Erkenntnis*, *88*(6), 2493–2514. https://doi.org/10.1007/s10670-021-00464-z
Hug, S. E., Ochsner, M., & Daniel, H.-D. (2013). Criteria for assessing research quality in the humanities: A Delphi study among scholars of English literature, German literature and art history. *Research Evaluation*, *22*(5), 369–383. https://doi.org/10.1093/reseval/rvt008
Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124
Ioannidis, J. P. A. (2012). Why Science Is Not Necessarily Self-Correcting. *Perspectives on Psychological Science*, *7*(6), 645–654. https://doi.org/10.1177/1745691612464056
Ioannidis, J. P. A. (2018). Meta-research: Why research on research matters. *PLOS Biology*, *16*(3), e2005468. https://doi.org/10.1371/journal.pbio.2005468
Iphofen, R. (2013). *Research Ethics in Ethnography/Anthropology*.
Israel, M. (2015). *Research Ethics and Integrity for Social Scientists: Beyond Regulatory Compliance*. SAGE Publications Ltd. https://doi.org/10.4135/9781473910096
Jasanoff, S. (1996). Beyond Epistemology: Relativism and Engagement in the Politics of Science. *Social Studies of Science*, *26*(2), 393–418. https://doi.org/10.1177/030631296026002008
Jasny, B. R., Chin, G., Chong, L., & Vignieri, S. (2011). Again, and Again, and Again …. *Science*, *334*(6060), 1225–1225. https://doi.org/10.1126/science.334.6060.1225
Kaiser, J., & Marshall, E. (1996). Imanishi-Kari Ruling Slams ORI: In finding that the Office of Research Integrity did not prove misconduct against a Tufts scientist, an HHS appeals panel also questioned how the government built its case. *Science*, *272*(5270), 1864–1866. https://doi.org/10.1126/science.272.5270.1864
Kaldewey, D. (2025). Die Krise der Faktizität, revisited: Wissenschaftsskepsis und Wissenschaftskommunikation im Kontext der COVID-19-Pandemie. *Österreichische Zeitschrift für Soziologie*, *50*(1), 22. https://doi.org/10.1007/s11614-025-00601-w
Kennedy, M.-R., Deans, Z., Ampollini, I., Breit, E., Bucchi, M., Seppel, K., Vie, K. J., & Meulen, R. T. (2023). "It is Very Difficult for us to Separate Ourselves from this System": Views of European Researchers, Research Managers, Administrators and Governance Advisors on Structural and Institutional Influences on Research Integrity. *Journal of Academic Ethics*, *21*(3), 471–495. https://doi.org/10.1007/s10805-022-09469-x
Kerkhoff, T., & Overeem, P. (2021). The Fluidity of Integrity: Lessons from Dutch Scandals. *Public Integrity*, *23*(1), 82–94. https://doi.org/10.1080/10999922.2020.1826139
Kerr, N. L. (1998). HARKing: Hypothesizing After the Results are Known. *Personality and Social Psychology Review*, *2*(3), 196–217. https://doi.org/10.1207/s15327957pspr0203_4

Kim, J., & Park, K. (2013). Ethical Modernization: Research Misconduct and Research Ethics Reforms in Korea Following the Hwang Affair. *Science and Engineering Ethics*, *19*(2), 355–380. https://doi.org/10.1007/s11948-011-9341-8

Kish-Gephart, J. J., Harrison, D. A., & Treviño, L. K. (2010). Bad apples, bad cases, and bad barrels: Meta-analytic evidence about sources of unethical decisions at work. *Journal of Applied Psychology, 95*(1), 1–31. https://doi.org/10.1037/a0017103

Klein, R. A., Cook, C. L., Ebersole, C. R., Vitiello, C., Nosek, B. A., Hilgard, J., Ahn, P. H., Brady, A. J., Chartier, C. R., Christopherson, C. D., Clay, S., Collisson, B., Crawford, J. T., Cromar, R., Gardiner, G., Gosnell, C. L., Grahe, J., Hall, C., Howard, I., … Ratliff, K. A. (2022). Many Labs 4: Failure to Replicate Mortality Salience Effect With and Without Original Author Involvement. *Collabra: Psychology*, *8*(1), 35271. https://doi.org/10.1525/collabra.35271

Klonsky, E. D. (2025). Campbell's Law Explains the Replication Crisis: Pre-Registration Badges Are History Repeating. *Assessment*, *32*(2), 224–234. https://doi.org/10.1177/10731911241253430

KNAW, NFU, NWO, TO2-Federatie, Vereniging Hogescholen, & VSNU. (2018). *Nederlandse gedragscode wetenschappelijke integriteit* [Application/pdf]. Data Archiving and Networked Services (DANS). https://doi.org/10.17026/DANS-2CJ-NVWU

Knorr, K. D. (1977). Producing and reproducing knowledge: Descriptive or constructive?: Toward a model of research production. *Social Science Information*, *16*(6), 669–696. https://doi.org/10.1177/053901847701600602

Knorr-Cetina, K. (1981). *The manufacture of knowledge: An essay on the constructivist and contextual nature of science*. Pergamon Press.

Knorr-Cetina, K. (1983). The ethnographer study of scientific work: Towards a constructivist
interpretation of science. In K. D. M. Knorr-Cetina (Eds.), Science observed: Perspectives on the social study of science. Bevery Hills, CA: Sage

Knorr-Cetina, K. (1995). Laboratory Studies: The Cultural Approach to the Study of Science. In S. Jasanoff, G. Markle, J. Peterson, & T. Pinch, *Handbook of Science and Technology Studies* (pp. 140–166). SAGE Publications, Inc. https://doi.org/10.4135/9781412990127.n7

Knorr-Cetina, K. (1999). *Epistemic cultures: How the sciences make knowledge*. Harvard University Press.

Knorr-Cetina, K. D. (1982). Scientific Communities or Transepistemic Arenas of Research? A Critique of Quasi-Economic Models of Science. *Social Studies of Science*, *12*(1), 101–130. https://doi.org/10.1177/030631282012001005

Koch, S. (1993). "Psychology" or "the psychological studies"? *American Psychologist*, *48*(8), 902–904. https://doi.org/10.1037/0003-066X.48.8.902

Kolstoe, S. E., & Pugh, J. (2024). The trinity of good research: Distinguishing between research integrity, ethics, and governance. *Accountability in Research*, *31*(8), 1222–1241. https://doi.org/10.1080/08989621.2023.2239712

König, Thomas, Vilain, Eric, & LoTempio, Jonathan Edward. (2024). *Open Science? Conceptualizing Openness as an Emerging Moral Economy of Science*. https://doi.org/10.3217/978-3-85125-976-6-02

Korbmacher, M., Azevedo, F., Pennington, C. R., Hartmann, H., Pownall, M., Schmidt, K., Elsherif, M., Breznau, N., Robertson, O., Kalandadze, T., Yu, S., Baker, B. J., O'Mahony, A., Olsnes, J. Ø.-S., Shaw, J. J., Gjoneska, B., Yamada, Y., Röer, J. P., Murphy, J., … Evans, T. (2023). The replication crisis has led to positive structural, procedural, and community changes. *Communications Psychology*, *1*(1), 3. https://doi.org/10.1038/s44271-023-00003-2

Kreimer, P. (2022). Constructivist Paradoxes Part 1: Critical Thoughts about Provincializing, Globalizing, and Localizing STS from a Non-Hegemonic Perspective. *Engaging Science, Technology, and Society*, *8*(2). https://doi.org/10.17351/ests2022.1109

Kretser, A., Murphy, D., Bertuzzi, S., Abraham, T., Allison, D. B., Boor, K. J., Dwyer, J., Grantham, A., Harris, L. J., Hollander, R., Jacobs-Young, C., Rovito, S., Vafiadis, D., Woteki, C., Wyndham, J., & Yada, R. (2019). Scientific Integrity Principles and Best Practices: Recommendations from a Scientific Integrity Consortium. *Science and Engineering Ethics*, *25*(2), 327–355. https://doi.org/10.1007/s11948-019-00094-3

Kuhn, T. S. (1962). *The structure of scientific revolutions*. University of Chicago Press. WorldCat.

Kulczycki, E. (2023). *The Evaluation Game: How Publication Metrics Shape Scholarly Communication* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009351218

Haven, T. L, & Van Grootel, Dr. L. (2019). Preregistering qualitative research. *Accountability in Research*, *26*(3), 229–244. https://doi.org/10.1080/08989621.2019.1580147

Labib, K., Crespo López, M. D. L. Á., Perello, C. P., & De Ridder, J. (2025). *Open Science as Confused: Contradictory and Conflicting Discourses in Open Science Guidance to Researchers*. https://doi.org/10.31222/osf.io/zr35u_v1

Labib, K., Roje, R., Bouter, L., Widdershoven, G., Evans, N., Marušić, A., Mokkink, L., & Tijdink, J. (2021). Important Topics for Fostering Research Integrity by Research Performing and Research Funding Organizations: A Delphi Consensus Study. *Science and Engineering Ethics*, *27*(4), 47. https://doi.org/10.1007/s11948-021-00322-9

Labib, K., Tijdink, J., Sijtsma, K., Bouter, L., Evans, N., & Widdershoven, G. (2024). How to combine rules and commitment in fostering research integrity? *Accountability in Research*, *31*(7), 917–943. https://doi.org/10.1080/08989621.2023.2191192

Laine, H. (2018). Open science and codes of conduct on research integrity. *Informaatiotutkimus*, *37*(4). https://doi.org/10.23978/inf.77414

Lakatos, I. (1970). Falsification and the Methodology of Scientific Research Programmes. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the Growth of*

Knowledge (1st ed., pp. 91–196). Cambridge University Press. https://doi.org/10.1017/CBO9781139171434.009
Lakens D. (2019). *The value of preregistration for psychological science: A conceptual analysis* (No. 3). 心理学評論刊行会. https://doi.org/10.24602/sjpr.62.3_221
Lakens, D. (2024). When and How to Deviate From a Preregistration. *Collabra: Psychology*, *10*(1), 117094. https://doi.org/10.1525/collabra.117094
Lakens, D., Mesquida, C., Rasti, S., & Ditroilo, M. (2024). The benefits of preregistration and Registered Reports. *Evidence-Based Toxicology*, *2*(1), 2376046. https://doi.org/10.1080/2833373X.2024.2376046
Lambert, M., & Degn, L. (2026). Shaping the Field: A Review of the Use of Theory in Research on Research Integrity. *Science and Engineering Ethics*, *32*(2), 19. https://doi.org/10.1007/s11948-026-00587-y
Lamont, M. (2009). *HOW PROFESSORS THINK*. Harvard University Press. JSTOR. http://www.jstor.org/stable/j.ctt13x0gx9
Lampland, M., & Star, S. L. (Eds.). (2009). *Standards and their stories: How quantifying, classifying, and formalizing practices shape everyday life*. Cornell University Press.
Langfeldt, L., Nedeva, M., Sörlin, S., & Thomas, D. A. (2020). Co-existing Notions of Research Quality: A Framework to Study Context-specific Understandings of Good Research. *Minerva*, *58*(1), 115–137. https://doi.org/10.1007/s11024-019-09385-2
Larregue, J. (2024). Sentencing social psychology: Scientific deviance and the diffusion of statistical rules. *Current Sociology*, *72*(3), 389–406. https://doi.org/10.1177/00113921221117604
Latour, B. (1987). *Science in action: How to follow scientists and engineers through society*. Harvard University Press.
Latour, B. (2004). Why has critique run out of steam? From matters of fact to matters of concern. *Critical Inquiry*, *30*(2), 225–248. https://doi.org/10.1086/421123
Latour, B., & Woolgar, S. (1979). *Laboratory life: The social construction of scientific facts*. Sage Publications.
Latour, B., & Woolgar, S. (1986). *Laboratory life: The construction of scientific facts* (2nd ed.). Princeton University Press
Laudel, G. (2024). Where do field-specific notions of research quality come from? *Research Evaluation*, *33*, rvae027. https://doi.org/10.1093/reseval/rvae027
Law, J. (2004). *After Method* (0 ed.). Routledge. https://doi.org/10.4324/9780203481141
LeBel, E. P. (2015). A New Replication Norm for Psychology. *Collabra*, *1*(1), 4. https://doi.org/10.1525/collabra.23
Leonelli, S. (2018). Rethinking Reproducibility as a Criterion for Research Quality. In L. Fiorito, S. Scheall, & C. E. Suprinyak (Eds.), *Research in the History of Economic Thought and Methodology* (Vol. 36, pp. 129–146). Emerald Publishing Limited. https://doi.org/10.1108/S0743-41542018000036B009

Leonelli, S. (2022). Open Science and Epistemic Diversity: Friends or Foes? *Philosophy of Science*, *89*(5), 991–1001. https://doi.org/10.1017/psa.2022.45
Leonelli, S. (2023). *Philosophy of Open Science* (1st ed.). Cambridge University Press. https://doi.org/10.1017/9781009416368
Li, D., & Cornelis, G. (2021). Differing perceptions concerning research misconduct between China and Flanders: A qualitative study. *Accountability in Research*, *28*(2), 63–94. https://doi.org/10.1080/08989621.2020.1802586
Lilienfeld, S. O. (2017). Psychology's Replication Crisis and the Grant Culture: Righting the Ship. *Perspectives on Psychological Science*, *12*(4), 660–664. https://doi.org/10.1177/1745691616687745
Lilja, E. (2021). Threat of policy alienation: Exploring the implementation of Open Science policy in research practice. *Science and Public Policy*, *47*(6), 803–817. https://doi.org/10.1093/scipol/scaa044
Lilja, E. (2025). Responsibility dynamics in open science: Exploring researchers' framing of ethical considerations in policy implementation. *Science and Public Policy*, *52*(6), 852–864. https://doi.org/10.1093/scipol/scaf021
Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. Sage Publications.
Lindemann, T., & Häberlein, L. (2023). Contours of a research ethics and integrity perspective on open science. *Frontiers in Research Metrics and Analytics*, *8*, 1052353. https://doi.org/10.3389/frma.2023.1052353
Li Vigni, F., Louvel, S., & Raimbault, B. (2023). How do the vectors of scientific credibility reconfigure at the interface between the social worlds? *Revue d'anthropologie Des Connaissances*, *17*(2). https://doi.org/10.4000/rac.30500
Lock, S. (1999). Book: The Baltimore Case: A Trial of Politics, Science, and Character. *BMJ*, *319*(7214), 926–926. https://doi.org/10.1136/bmj.319.7214.926
Lubega, N., Anderson, A., & Nelson, N. C. (2023). Experience of irreproducibility as a risk factor for poor mental health in biomedical science doctoral students: A survey and interview-based study. *PLOS ONE*, *18*(11), e0293584. https://doi.org/10.1371/journal.pone.0293584
Lund, A. (2025). An alternative open science framework. *Journal of Documentation*, *81*(7), 157–178. https://doi.org/10.1108/JD-09-2024-0220
Lynch, M. (1982). TECHNICAL WORK AND CRITICAL INQUIRY - INVESTIGATIONS IN A SCIENTIFIC LABORATORY. *SOCIAL STUDIES OF SCIENCE*, *12*(4), 499–533. (WOS:A1982PU59500002). https://doi.org/10.1177/030631282012004002
Lynch, M. (1985). *Art and artifact in laboratory science: A study of shop work and shop talk in a research laboratory*. WorldCat.
Lynch, M. (2020). We Have Never Been Anti-Science: Reflections on Science Wars and Post-Truth. *Engaging Science, Technology, and Society*, *6*, 49–57. https://doi.org/10.17351/ests2020.309
Lynch, M. (2024). Laboratory studies and scientific practice. In U. Felt & A. Irwin (Eds.), *Elgar Encyclopedia of Science and Technology Studies* (pp. 273–282). Edward Elgar Publishing. https://doi.org/10.4337/9781800377998.ch28

Lynch, M., Gerst, D., Krämer, H., & Salomon, R. (2019). The Studies are Probably the Best Thing That Garfinkel Ever Wrote. Michael Lynch in Conversation With Dominik Gerst, Hannes Krämer & René Salomon. *Forum Qualitative Sozialforschung Forum: Qualitative Social Research*, *20*(2). https://doi.org/10.17169/fqs-20.2.3251

Ma, L. (2025). The limits of openness: Global knowledge production and its boundaries. *Journal of Documentation*, *81*(7), 121–134. https://doi.org/10.1108/JD-09-2024-0237

Ma, L., & Davin, C. (2025). *For Researchers in the Humanities, Is Open Really Fair?* https://doi.org/10.1146/katina-111825-1

Macfarlane, B. (2024). The DECAY of Merton's scientific norms and the new academic ethos. *Oxford Review of Education*, *50*(4), 468–483. https://doi.org/10.1080/03054985.2023.2243814

Macleod, M. (2021). Want research integrity? Stop the blame game. *Nature*, *599*(7886), 533–533. https://doi.org/10.1038/d41586-021-03493-4

Macleod, M. (2022). Improving the reproducibility and integrity of research: What can different stakeholders contribute? *BMC Research Notes*, *15*(1), 146. https://doi.org/10.1186/s13104-022-06030-2

Makel, M. C., & Plucker, J. A. (2014). Facts Are More Important Than Novelty: Replication in the Education Sciences. *Educational Researcher*, *43*(6), 304–316. https://doi.org/10.3102/0013189X14545513

Makel, M. C., Plucker, J. A., & Hegarty, B. (2012). Replications in Psychology Research: How Often Do They Really Occur? *Perspectives on Psychological Science*, *7*(6), 537–542. https://doi.org/10.1177/1745691612460688

Malich, L., & Rehmann-Sutter, C. (2022). Metascience Is Not Enough – A Plea for Psychological Humanities in the Wake of the Replication Crisis. *Review of General Psychology*, *26*(2), 261–273. https://doi.org/10.1177/10892680221083876

Malički, M., Katavić, V., Marković, D., Marušić, M., & Marušić, A. (2019). Perceptions of Ethical Climate and Research Pressures in Different Faculties of a University: Cross-Sectional Study at the University of Split, Croatia. *Science and Engineering Ethics*, *25*(1), 231–245. https://doi.org/10.1007/s11948-017-9987-y

Manco, A. (2022). A Landscape of Open Science Policies Research. *Sage Open*, *12*(4), 21582440221140358. https://doi.org/10.1177/21582440221140358

Manco, A. (2024). Open science policies as regarded by the communities of researchers from the basic sciences in the scientific periphery. *Online Information Review*, *48*(6), 1065–1087. https://doi.org/10.1108/OIR-03-2023-0135

Manco, A. (2025). *Exploring Open Science and Research Integrity in the Basic Sciences: Findings from a Qualitative Study*. MetaArXiv. https://doi.org/10.31222/osf.io/98yds_v2

Matarese, V. (2022). Kinds of Replicability: Different Terms and Different Functions. *Axiomathes*, *32*(S2), 647–670. https://doi.org/10.1007/s10516-021-09610-2

Matarese, V. (2023). A new concept of replication. *Inquiry*, 1–26. https://doi.org/10.1080/0020174X.2023.2278032

McDermott, R. (2022). Breaking free: How preregistration hurts scholars and science. *Politics and the Life Sciences*, *41*(1), 55–59. https://doi.org/10.1017/pls.2022.4

Medawar, B. (1963). *Is the scientific paper a fraud? In the threat and the glory: Reflections on science and scientists*. New York: Harper Collins.

Meirmans, S., Luijken, K., Bruijning-Verhagen, P. C. J. L., Groenwold, R. H. H., & de Boer, A. R. (2025). The function of conducting a direct replication may be different than typically assumed: Understanding contextuality and asymmetry. *Journal of Clinical Epidemiology*, *183*. https://doi.org/10.1016/j.jclinepi.2025.111785

Mejlgaard, N., Bouter, L. M., Gaskell, G., Kavouras, P., Allum, N., Bendtsen, A.-K., Charitidis, C. A., Claesen, N., Dierickx, K., Domaradzka, A., Reyes Elizondo, A., Foeger, N., Hiney, M., Kaltenbrunner, W., Labib, K., Marušić, A., Sørensen, M. P., Ravn, T., Ščepanović, R., … Veltri, G. A. (2020). Research integrity: Nine ways to move from talk to walk. *Nature*, *586*(7829), 358–360. https://doi.org/10.1038/d41586-020-02847-8

Merton, R. K. (1942). A note on science and democracy. Journal of Legal and Political Sociology, 1(1), 15–26

Merton, R. K. (1957). Priorities in Scientific Discovery: A Chapter in the Sociology of Science. *American Sociological Review*, *22*(6), 635. https://doi.org/10.2307/2089193

Merton, R. K. (1963). The ambivalence of scientists. *Bulletin of the Johns Hopkins Hospital*, *112*, 77–97.

Merton, R. K. (1973). *The sociology of science: Theoretical and empirical investigations*. University of Chicago press.

Merton, R. K., & Barber, E. (1963). Sociological ambivalence. *Sociological Ambivalence and Other Essays*, 3-31.

Meskus, M., Marelli, L., & D'Agostino, G. (2018). Research Misconduct in the Age of Open Science: The Case of STAP Stem Cells. *Science as Culture*, *27*(1), 1–23. https://doi.org/10.1080/09505431.2017.1316975

Mickel, A., Heidebrink-Bruno, A., Hosseiny Marani, A., Barone, I., Lee, O., & Baumer, E. P. S. (2023). The Cultural Production of Everyday Ethics in Two University STEM Labs. *Bulletin of Science, Technology & Society*, *43*(1–2), 3–17. https://doi.org/10.1177/02704676231181030

Millerand, F., Ribes, D., Baker, K. S., & Bowker, G. C. (2013). Making an Issue out of a Standard: Storytelling Practices in a Scientific Community. *Science,*

Technology, & Human Values, *38*(1), 7–43. https://doi.org/10.1177/0162243912437221
Mitroff, I. I. (1974). Norms and Counter-Norms in a Select Group of the Apollo Moon Scientists: A Case Study of the Ambivalence of Scientists. *American Sociological Review*, *39*(4), 579. https://doi.org/10.2307/2094423
Mol, A., & Law, J. (1994). Regions, Networks and Fluids: Anaemia and Social Topology. *Social Studies of Science*, *24*(4), 641–671. JSTOR.
Moon, K., & Blackman, D. (2014). A Guide to Understanding Social Science Research for Natural Scientists. *Conservation Biology*, *28*(5), 1167–1177. https://doi.org/10.1111/cobi.12326
Moors, A. (2026). *Doing Ethnography*. Leuven University Press. https://doi.org/10.11116/9789461667502
Morawski, J. (2020). Psychologists' psychologies of psychologists in a time of crisis. *History of Psychology*, *23*(2), 176–198. https://doi.org/10.1037/hop0000140
Morris, S. A., & Van Der Veer Martens, B. (2008). Mapping research specialties. *Annual Review of Information Science and Technology*, *42*(1), 213–295. https://doi.org/10.1002/aris.2008.1440420113
Moskovkin, V. (2024, July 15). Tracing the origins of 'publish or perish'—LSE Impact. *LSE Impact - Understanding Impact and Practice in Academic Research*. https://blogs.lse.ac.uk/impactofsocialsciences/2024/07/15/tracing-the-origins-of-publish-or-perish/
Mulkay, M. (1984). The Scientist Talks Back: A One-Act Play, with a Moral, about Replication in Science and Reflexivity in Sociology. *Social Studies of Science*, *14*(2), 265–282. https://doi.org/10.1177/030631284014002008
Mulkay, M., & Gilbert, G. N. (1986). Replication and Mere Replication. *Philosophy of the Social Sciences*, *16*(1), 21–37. https://doi.org/10.1177/004839318601600102
Mulkay, M. J. (1976). Norms and ideology in science. *Social Science Information*, *15*(4–5), 637–656. https://doi.org/10.1177/053901847601500406
Müller, K. (2025). *Quality and the Humanities: Negotiating Research Quality in Swedish Humanities, 1980s–2020s*. KTH Royal Institute of Technology.
Müller, K., Salö, L., & Sörlin, S. (2024). Quality from within: Entry points to research quality in the humanities. *Research Evaluation*, *33*, rvae029. https://doi.org/10.1093/reseval/rvae029
Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., Percie Du Sert, N., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., & Ioannidis, J. P. A. (2017). A manifesto for reproducible science. *Nature Human Behaviour*, *1*(1), 0021. https://doi.org/10.1038/s41562-016-0021
Murphy, A. K., Jerolmack, C., & Smith, D. (2021). Ethnography, Data Transparency, and the Information Age. *Annual Review of Sociology*, *47*(1), 41–61. https://doi.org/10.1146/annurev-soc-090320-124805
Muthanna, A., Chaaban, Y., & Qadhi, S. (2024). A model of the interrelationship between research ethics and research integrity. *International Journal of*

Qualitative Studies on Health and Well-Being, *19*(1), 2295151. https://doi.org/10.1080/17482631.2023.2295151
Naddaf, M. (2025). Research-integrity sleuths say their work is being 'twisted' to undermine science. *Nature*, *643*(8072), 615–616. https://doi.org/10.1038/d41586-025-02163-z
Naeem, M., Ozuem, W., Howell, K., & Ranfagni, S. (2024). Demystification and Actualisation of Data Saturation in Qualitative Research Through Thematic Analysis. *International Journal of Qualitative Methods*, *23*, 16094069241229777. https://doi.org/10.1177/16094069241229777
Nagai, H., Nakazawa, E., & Akabayashi, A. (2022). The creation of the Belmont Report and its effect on ethical principles: A historical study. *Monash Bioethics Review*, *40*(2), 157–170. https://doi.org/10.1007/s40592-022-00165-5
National Academy of Sciences (US), National Academy of Engineering (US) and Institute of Medicine (US) Panel on Scientific Responsibility and the Conduct of Research. (1992). *Responsible Science: Ensuring the Integrity of the Research Process*. National Academies Press (US). https://doi.org/10.17226/1864
National Academies of Sciences, Engineering, and Medicine. (2017). *Fostering integrity in research*. The National Academies Press.
Nelson, L. D., Simmons, J., & Simonsohn, U. (2018). Psychology's Renaissance. *Annual Review of Psychology*, *69*(1), 511–534. https://doi.org/10.1146/annurev-psych-122216-011836
Nelson, N. C. (2021). Understand the real reasons reproducibility reform fails. *Nature*, *600*(7888), 191–191. https://doi.org/10.1038/d41586-021-03617-w
Nelson, N. C., Chung, J., Ichikawa, K., & Malik, M. M. (2022). Psychology Exceptionalism and the Multiple Discovery of the Replication Crisis. *Review of General Psychology*, *26*(2), 184–198. https://doi.org/10.1177/10892680211046508
Nelson, N. C., Ichikawa, K., Chung, J., & Malik, M. M. (2021). Mapping the discursive dimensions of the reproducibility crisis: A mixed methods analysis. *PLOS ONE*, *16*(7), e0254090. https://doi.org/10.1371/journal.pone.0254090
Neoh, M. J. Y., Carollo, A., Lee, A., & Esposito, G. (2023). Fifty years of research on questionable research practises in science: Quantitative analysis of co-citation patterns. *Royal Society Open Science*, *10*(10), 230677. https://doi.org/10.1098/rsos.230677
Nguyen, C. T. (2022). Transparency is Surveillance. *Philosophy and Phenomenological Research*, *105*(2), 331–361. https://doi.org/10.1111/phpr.12823
Nguyen, L. T., & Tuamsuk, K. (2024). Unveiling scientific integrity in scholarly publications: A bibliometric approach. *International Journal for Educational Integrity*, *20*(1), 16. https://doi.org/10.1007/s40979-024-00164-5
Norton, J. D. (2015). Replicability of Experiment. *Theoria: Revista de Teoría, Historia y Fundamentos de La Ciencia*, *30*(2), 229.

Nosek, B. (2019, June 11). *Strategy for Culture Change*. https://www.cos.io/blog/strategy-for-culture-change
Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J., Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., … Yarkoni, T. (2015). Promoting an open research culture. *Science*, *348*(6242), 1422–1425. https://doi.org/10.1126/science.aab2374
Nosek, B. A., Beck, E. D., Campbell, L., Flake, J. K., Hardwicke, T. E., Mellor, D. T., Van 'T Veer, A. E., & Vazire, S. (2019). Preregistration Is Hard, And Worthwhile. *Trends in Cognitive Sciences*, *23*(10), 815–818. https://doi.org/10.1016/j.tics.2019.07.009
Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, *115*(11), 2600–2606. https://doi.org/10.1073/pnas.1708274114
Nosek, B. A., & Errington, T. M. (2017). Making sense of replications. *eLife*, *6*, e23383. https://doi.org/10.7554/eLife.23383
Nosek, B. A., & Errington, T. M. (2020). What is replication? *PLOS Biology*, *18*(3), e3000691. https://doi.org/10.1371/journal.pbio.3000691
Nosek, B. A., & Lindsay, S. D. (2018). Preregistration Becoming the Norm in Psychological Science. *APS Observer*, *31*. https://www.psychologicalscience.org/observer/preregistration-becoming-the-norm-in-psychological-science
Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific Utopia: II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science*, *7*(6), 615–631. https://doi.org/10.1177/1745691612459058
Nuzzo, R. (2015). How scientists fool themselves – and how they can stop. *Nature*, *526*(7572), 182–185. https://doi.org/10.1038/526182a
Oberauer, K., & Lewandowsky, S. (2019). Addressing the theory crisis in psychology. *Psychonomic Bulletin & Review*, *26*(5), 1596–1618. https://doi.org/10.3758/s13423-019-01645-2
Ochsner, M. (2022). Identifying research quality in the social sciences. In T. C. E. Engels & E. Kulczycki (Eds.), *Handbook on research assessment in the social sciences* (pp. 48–66). Edward Elgar Publishing. https://doi.org/10.4337/9781800372559
Office of Science and Technology Policy. (2000, December 6). Federal research misconduct policy (65 FR 76260). Federal Register. https://www.federalregister.gov/documents/2000/12/06/00-30852/executive-office-of-the-president-federal-policy-on-research-misconduct-preamble-for-research
Ofosu, G. K., & Posner, D. N. (2023). Pre-Analysis Plans: An Early Stocktaking. *Perspectives on Politics*, *21*(1), 174–190. https://doi.org/10.1017/S1537592721000931

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*(6251), aac4716. https://doi.org/10.1126/science.aac4716
Orkeny, B. (2025). *The Epistemic Function of Replication: Mapping the Domains of the Debate*. Open Science Framework. https://doi.org/10.31219/osf.io/h8swj_v1
Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, *17*(11), 776–783. https://doi.org/10.1037/h0043424
Ottolini, L., & Noel, M. (2024, June 26). Opening up science as a work: An international comparison of openness to society and openness of publication. *NECSUS*. https://necsus-ejms.org/opening-up-science-as-a-work-an-international-comparison-of-openness-to-society-and-openness-of-publication/
Pachirat, T. (2018). *Among wolves: Ethnography and the immersive study of power*. Routledge. https://doi.org/10.4324/9780203701102
Packer, K., & Webster, A. (1996). Patenting Culture in Science: Reinventing the Scientific Wheel of Credibility. *Science, Technology, & Human Values*, *21*(4), 427–453. JSTOR.
Panel on Scientific Responsibility and the Conduct of Research, Committee on Science, Engineering, and Public Policy, National Academy of Sciences, National Academy of Engineering, & Institute of Medicine. (1992). *Responsible Science: Ensuring the Integrity of the Research Process: Volume I* (p. 1864). National Academies Press. https://doi.org/10.17226/1864
Paradeise, C., & Filliatreau, G. (2021). Scientific Integrity Matters. *Minerva*, *59*(3), 289–309. https://doi.org/10.1007/s11024-021-09440-x
Pascal, C. B. (1999). The history and future of the office of research integrity: Scientific misconduct and beyond. *Science and Engineering Ethics*, *5*(2), 183–198. https://doi.org/10.1007/s11948-999-0008-7
Pashler, H., & Harris, C. R. (2012). Is the Replicability Crisis Overblown? Three Arguments Examined. *Perspectives on Psychological Science*, *7*(6), 531–536. https://doi.org/10.1177/1745691612463401
Patton, C. (2024). Replicability and the humanities: The problem with universal measures of research quality. *Research Evaluation*, *34*, rvaf052. https://doi.org/10.1093/reseval/rvaf052
Peels, R. (2019). Replicability and replication in the humanities. *Research Integrity and Peer Review*, *4*(1), 2. https://doi.org/10.1186/s41073-018-0060-4
Peels, R., & Bouter, L. (2023). Replication and trustworthiness. *Accountability in Research*, *30*(2), 77–87. https://doi.org/10.1080/08989621.2021.1963708
Peels, R., De Ridder, J., Haven, T., & Bouter, L. (2019). Value pluralism in research integrity. *Research Integrity and Peer Review*, *4*(1), 18. https://doi.org/10.1186/s41073-019-0076-4
Penders, B. (2022). Process and Bureaucracy: Scientific Reform as Civilisation. *Bulletin of Science, Technology & Society*, *42*(4), 107–116. https://doi.org/10.1177/02704676221126388

Penders, B. (2024). Scandal in scientific reform: The breaking and remaking of science. *Journal of Responsible Innovation*, *11*(1), 2371172. https://doi.org/10.1080/23299460.2024.2371172
Penders, B., De Rijcke, S., & Holbrook, J. B. (2020). Science's moral economy of repair: Replication and the circulation of reference. *Accountability in Research*, *27*(2), 107–113. https://doi.org/10.1080/08989621.2020.1720659
Penders, B., Holbrook, J. B., & De Rijcke, S. (2019). Rinse and Repeat: Understanding the Value of Replication across Different Ways of Knowing. *Publications*, *7*(3), 52. https://doi.org/10.3390/publications7030052
Penders, B., & Janssens, A. C. J. W. (2018). Finding Wealth in Waste: Irreplicability Re-Examined. *BioEssays*, *40*(12), 1800173. https://doi.org/10.1002/bies.201800173
Penders, B., & Nelis, A. P. (2011). Credibility Engineering in the Food Industry: Linking Science, Regulation, and Marketing in a Corporate Context. *Science in Context*, *24*(4), 487–515. https://doi.org/10.1017/S0269889711000202
Penders, B., Vos, R., & Horstman, K. (2009). A question of style: Method, integrity and the meaning of proper science. *Endeavour*, *33*(3), 93–98. https://doi.org/10.1016/j.endeavour.2009.07.001
Peng, R. D., Dominici, F., & Zeger, S. L. (2006). Reproducible Epidemiologic Research. *American Journal of Epidemiology*, *163*(9), 783–789. https://doi.org/10.1093/aje/kwj093
Peršak, N. (2025). Research Integrity and Qualitative Research: Tackling and Researching Unethical Practices at the Top. In H. Busby (Ed.), *Advances in Research Ethics and Integrity* (pp. 235–251). Emerald Publishing Limited. https://doi.org/10.1108/S2398-601820250000012014
Peruginelli, G., & Pölönen, J. (2024). The legal foundation of responsible research assessment: An overview on European Union and Italy. *Research Evaluation*, *32*(4), 670–682. https://doi.org/10.1093/reseval/rvad035
Peterson, D. (2016). The Baby Factory: Difficult Research Objects, Disciplinary Standards, and the Production of Statistical Significance. *Socius: Sociological Research for a Dynamic World*, *2*, 2378023115625071. https://doi.org/10.1177/2378023115625071
Peterson, D. (2021). The replication crisis won't be solved with broad brushstrokes. *Nature*, *594*(7862), 151–151. https://doi.org/10.1038/d41586-021-01509-7
Peterson, D., & Panofsky, A. (2021). Self-correction in science: The diagnostic and integrative motives for replication. *Social Studies of Science*, *51*(4), 583–605. https://doi.org/10.1177/03063127211005551
Peterson, D., & Panofsky, A. (2023). Metascience as a Scientific Social Movement. *Minerva*, *61*(2), 147–174. https://doi.org/10.1007/s11024-023-09490-3
Pham, M. T., & Oh, T. T. (2021). Preregistration Is Neither Sufficient nor Necessary for Good Science. *Journal of Consumer Psychology*, *31*(1), 163–176. https://doi.org/10.1002/jcpy.1209

Philipps, A., & Paruschke, L. (2024). Inside regular lab meetings: The social construction of a research team and ideas in optical physics. *Social Studies of Science*, *54*(2), 257–280. https://doi.org/10.1177/03063127231188132
Pickering, A. (1992). From science as knowledge to science as practice. In A. Pickering (Ed.), *Science as practice and culture* (p. 4). University of Chicago Press.
Pickersgill, M. (2012). The Co-production of Science, Ethics, and Emotion. *Science, Technology, & Human Values*, *37*(6), 579–603. https://doi.org/10.1177/0162243911433057
Pickersgill, M. (2021). Experiences of Ethics, Governance and Scientific Practice in Neuroscience Research. In G. Laurie, E. Dove, A. Ganguli-Mitra, C. McMillan, E. Postan, N. Sethi, & A. Sorbie (Eds.), *The Cambridge Handbook of Health Research Regulation* (1st ed., pp. 306–314). Cambridge University Press. https://doi.org/10.1017/9781108620024.038
Plesser, H. E. (2018). Reproducibility vs. Replicability: A Brief History of a Confused Terminology. *Frontiers in Neuroinformatics*, *11*, 76. https://doi.org/10.3389/fninf.2017.00076
Polanyi, M. (1958). *Personal Knowledge: Towards a Post-Critical Philosophy* (M. J. Nye, Ed.). University of Chicago Press.
Pols, J., M'charek, A., Jerak-Zuiderent, S., & Brenninkmeijer, J. (2024). Achieving good science: The integrity of scientific institutions. *Learning and Teaching*, *17*(1), 24–53. https://doi.org/10.3167/latiss.2024.170103
Ponterotto, J. (2015). Brief Note on the Origins, Evolution, and Meaning of the Qualitative Research Concept Thick Description. *The Qualitative Report*. https://doi.org/10.46743/2160-3715/2006.1666
Popper, K. R. (1959). *The logic of scientific discovery*. Basic Books.
Potthoff, S., Hempeler, C., Gather, J., Gieselmann, A., Vollmann, J., & Scholten, M. (2023). Research ethics in practice: An analysis of ethical issues encountered in qualitative health research with mental health service users and relatives. *Medicine, Health Care and Philosophy*, *26*(4), 517–527. https://doi.org/10.1007/s11019-023-10169-5
Price, A. R. (2013). Research Misconduct and Its Federal Regulation: The Origin and History of the Office of Research Integrity—With Personal Views by ORI's Former Associate Director for Investigative Oversight. *Accountability in Research*, *20*(5–6), 291–319. https://doi.org/10.1080/08989621.2013.822238
Price, D. J. de S. (1963). *Little Science, Big Science*. Columbia University Press. https://doi.org/10.7312/pric91844
Prosser, A. M., Bagnall, R., & Higson-Sweeney, N. (2024). Reflection over compliance: Critiquing mandatory data sharing policies for qualitative research. *Journal of Health Psychology*, *29*(7), 653–658. https://doi.org/10.1177/13591053231225903
Protzko, J., Krosnick, J., Nelson, L., Nosek, B. A., Axt, J., Berent, M., Buttrick, N., DeBell, M., Ebersole, C. R., Lundmark, S., MacInnis, B., O'Donnell, M., Perfecto, H., Pustejovsky, J. E., Roeder, S. S., Walleczek, J., & Schooler, J. W.

(2023). RETRACTED ARTICLE: High replicability of newly discovered social-behavioural findings is achievable. *Nature Human Behaviour*, *8*(2), 311–319. https://doi.org/10.1038/s41562-023-01749-9

Protzko, J., Krosnick, J., Nelson, L., Nosek, B. A., Axt, J., Berent, M., Buttrick, N., DeBell, M., Ebersole, C. R., Lundmark, S., MacInnis, B., O'Donnell, M., Perfecto, H., Pustejovsky, J. E., Roeder, S. S., Walleczek, J., & Schooler, J. W. (2024). Retraction Note: High replicability of newly discovered social-behavioural findings is achievable. *Nature Human Behaviour*, *8*(10), 2067–2067. https://doi.org/10.1038/s41562-024-01997-3

Protzko, J., & Schooler, J. W. (2020). No relationship between researcher impact and replication effect: An analysis of five studies with 100 replications. *PeerJ*, *8*, e8014. https://doi.org/10.7717/peerj.8014

Pulverer, B. (2014). STAP dance. *The EMBO Journal*, *33*(12), 1285–1286. https://doi.org/10.15252/embj.201489076

Ramsey, R. (2021). A Call for Greater Modesty in Psychology and Cognitive Neuroscience. *Collabra: Psychology*, *7*(1), 24091. https://doi.org/10.1525/collabra.24091

Raven, P. H., & Marks, J. (2009). Scientific Misconduct. In *Why I Am Not a Scientist* (1st ed., pp. 162–197). University of California Press. JSTOR. http://www.jstor.org/stable/10.1525/j.ctt1pnstm.10

Ravn, T., & Sørensen, M. P. (2021). Exploring the Gray Area: Similarities and Differences in Questionable Research Practices (QRPs) Across Main Areas of Research. *Science and Engineering Ethics*, *27*(4), 40. https://doi.org/10.1007/s11948-021-00310-z

Redman, B. K., Yarandi, H. N., & Merz, J. F. (2008). Empirical developments in retraction. *Journal of Medical Ethics*, *34*(11), 807–809. https://doi.org/10.1136/jme.2007.023069

Reinhart, M. (2016). Reproducibility in the Social Sciences. In H. Atmanspacher & S. Maasen (Eds.), *Reproducibility: Principles, Problems, Practices, and Prospects* (pp. 407–423). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118865064.ch19

Reischer, H. N., & Cowan, H. R. (2020). Quantity Over Quality? Reproducible Psychological Science from a Mixed Methods Perspective. *Collabra: Psychology*, *6*(1), 26. https://doi.org/10.1525/collabra.284

Resnik, D. B. (2019). Is it time to revise the definition of research misconduct? *Accountability in Research*, *26*(2), 123–137. https://doi.org/10.1080/08989621.2019.1570156

Resnik, D. B., Lee, E., Jirles, B., Smith, E., & Barker, K. (2023). For the "good of the lab": Insights from three focus groups concerning the ethics of managing a laboratory or research group. *Accountability in Research*, *30*(4), 199–218. https://doi.org/10.1080/08989621.2021.1983799

Resnik, D. B., Neal, T., Raymond, A., & Kissling, G. E. (2015). Research Misconduct Definitions Adopted by U.S. Research Institutions. *Accountability in Research*, *22*(1), 14–21. https://doi.org/10.1080/08989621.2014.891943

Resnik, D. B., & Shamoo, A. E. (2017). Reproducibility and Research Integrity. *Accountability in Research*, *24*(2), 116–123. https://doi.org/10.1080/08989621.2016.1257387
Resnik, D. B., & Stewart, C. N. (2012). Misconduct versus honest error and scientific disagreement. *Accountability in Research*, *19*(1), 56–63. https://doi.org/10.1080/08989621.2012.650948
Resnik, D. B., Stewart Jr, C. N., Williams, F., Thiele, C., Yamada, K. M., & Barker, K. (2024). Ethical decision-making and role conflict in managing a scientific laboratory. *Accountability in Research*, *31*(8), 1198–1221. https://doi.org/10.1080/08989621.2023.2236553
Rip, A. (1994). The republic of science in the 1990s. *Higher Education*, *28*(1), 3–23. https://doi.org/10.1007/BF01383569
Ritchie, S. J., Wiseman, R., & French, C. C. (2012). Failing the Future: Three Unsuccessful Attempts to Replicate Bem's 'Retroactive Facilitation of Recall' Effect. *PLoS ONE*, *7*(3), e33423. https://doi.org/10.1371/journal.pone.0033423
Roberts, B. W., Jackson, J. J., Fayard, J. V., Edmonds, G., & Meints, J. (2009). Conscientiousness. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 369–381). The Guilford Press.
Roberts, L. L., Sibum, H. O., & Mody, C. C. M. (2020). Integrating the history of science into broader discussions of research integrity and fraud. *History of Science*, *58*(4), 354–368. https://doi.org/10.1177/0073275320952268
Robinson-Garcia, N., Costas, R., Nane, G. F., & Van Leeuwen, T. N. (2023). Valuation regimes in academia: Researchers' attitudes towards their diversity of activities and academic performance. *Research Evaluation*, *32*(2), 496–514. https://doi.org/10.1093/reseval/rvac049
Romero, F. (2018). Who should do replication labor? *Advances in Methods and Practices in Psychological Science, 1*(4), 516–537. https://doi.org/10.1177/2515245918803619
Romero, F. (2019). Philosophy of science and the replicability crisis. *Philosophy Compass*, *14*(11), e12633.
Rosenthal, R. (1979). The file drawer problem and tolerance for null results. *Psychological Bulletin*, *86*(3), 638–641. https://doi.org/10.1037/0033-2909.86.3.638
Ross-Hellauer, T., Klebel, T., Bannach-Brown, A., Horbach, S. P. J. M., Jabeen, H., Manola, N., Metodiev, T., Papageorgiou, H., Reczko, M., Sansone, S.-A., Schneider, J., Tijdink, J., & Vergoulis, T. (2022). TIER2: Enhancing Trust, Integrity and Efficiency in Research through next-level Reproducibility. *Research Ideas and Outcomes*, *8*, e98457. https://doi.org/10.3897/rio.8.e98457
Rubin, M. (2022). The Costs of HARKing. *The British Journal for the Philosophy of Science*, *73*(2), 535–560. https://doi.org/10.1093/bjps/axz050
Rubin, M. (2024). Questionable Metascience Practices. *Journal of Trial and Error*, *4*(1), 5–20. https://doi.org/10.36850/mr4

Rubin, M., & Donkin, C. (2024). Exploratory hypothesis tests can be more compelling than confirmatory hypothesis tests. *Philosophical Psychology*, *37*(8), 2019–2047. https://doi.org/10.1080/09515089.2022.2113771

Rushforth, A., & Hammarfelt, B. (2023). The rise of responsible metrics as a professional reform movement: A collective action frames account. *Quantitative Science Studies*, *4*(4), 879–897. https://doi.org/10.1162/qss_a_00280

Sandler, J., & Thedvall, R. (2017). *Meeting Ethnography: Meetings as Key Technologies of Contemporary Governance, Development, and Resistance* (1st ed.). Routledge. https://doi.org/10.4324/9781315559407

Sarafoglou, A., Kovacs, M., Bakos, B., Wagenmakers, E.-J., & Aczel, B. (2022). A survey on how preregistration affects the research workflow: Better science but more work. *Royal Society Open Science*, *9*(7), 211997. https://doi.org/10.1098/rsos.211997

Sarisoy, J. (2025). Why We Disagree About the Success of Replications. *Journal for General Philosophy of Science*, *56*(3), 307–324. https://doi.org/10.1007/s10838-024-09709-1

Saunders, B., Kitzinger, J., & Kitzinger, C. (2015). Anonymising interview data: Challenges and compromise in practice. *Qualitative Research*, *15*(5), 616–632. https://doi.org/10.1177/1468794114550439

Saunders, B., Sim, J., Kingstone, T., Baker, S., Waterfield, J., Bartlam, B., Burroughs, H., & Jinks, C. (2018). Saturation in qualitative research: Exploring its conceptualization and operationalization. *Quality & Quantity*, *52*(4), 1893–1907. https://doi.org/10.1007/s11135-017-0574-8

Schickore, J., & Hangel, N. (2019). "It might be this, it should be that…" uncertainty and doubt in day-to-day research practice. *European Journal for Philosophy of Science*, *9*(2), 31. https://doi.org/10.1007/s13194-019-0253-9

Schmidt, M. (2024). Why do some retracted articles continue to get cited? *Scientometrics*, *129*(12), 7535–7563. https://doi.org/10.1007/s11192-024-05147-4

Schmidt, S. (2009). Shall we really do it again? The powerful concept of replication is neglected in the social sciences. *Review of General Psychology*, *13*(2), 90–100. https://doi.org/10.1037/a0015108

Schneider, B., Ehrhart, M. G., & Macey, W. H. (2013). Organizational Climate and Culture. *Annual Review of Psychology*, *64*(1), 361–388. https://doi.org/10.1146/annurev-psych-113011-143809

Schneider, J. W., Allum, N., Andersen, J. P., Petersen, M. B., Madsen, E. B., Mejlgaard, N., & Zachariae, R. (2024). Is something rotten in the state of Denmark? Cross-national evidence for widespread involvement but not systematic use of questionable research practices across all fields of research. *PLOS ONE*, *19*(8), e0304342. https://doi.org/10.1371/journal.pone.0304342

Schneider, J. W., Horbach, S. P. J. M., & Aagaard, K. (2021). Stop blaming external factors: A historical-sociological argument. *Social Science Information*, *60*(3), 329–337. https://doi.org/10.1177/05390184211018123

Schniedermann, A. (2025). *From Form to Formation. Biomedical Reporting Standards in Practice*. https://doi.org/10.31222/osf.io/b8439_v2
Shapin, S. (1984). Pump and Circumstance: Robert Boyle's Literary Technology. *Social Studies of Science*, *14*(4), 481–520. JSTOR.
Shapin, S. (1988). Understanding the Merton Thesis. *Isis*, *79*(4), 594–605. https://doi.org/10.1086/354847
Shapin, S. (1995a). Cordelia's Love: Credibility and the Social Studies of Science. *Perspectives on Science*, *3*(3), 255–275. https://doi.org/10.1162/posc_a_00484
Shapin, S. (1995b). Here and everywhere: Sociology of scientific knowledge. *Annual Review of Sociology*, *21*(1), 289–321. https://doi.org/10.1146/annurev.so.21.080195.001445
Shapin, S., & Schaffer, S. (1985). *Leviathan and the Air-Pump: Hobbes, Boyle, and the Experimental Life* (REV-Revised). Princeton University Press. https://doi.org/10.2307/j.ctt7sv46
Shaw, D. (2019). The Quest for Clarity in Research Integrity: A Conceptual Schema. *Science and Engineering Ethics*, *25*(4), 1085–1093. https://doi.org/10.1007/s11948-018-0052-2
Shaw, D., & Satalkar, P. (2018). Researchers' interpretations of research integrity: A qualitative study. *Accountability in Research*, *25*(2), 79–93. https://doi.org/10.1080/08989621.2017.1413940
Shore, C. (2024). The bureaucratization of ethical integrity: Research ethics committees and imaginaries of risk. *Anthropology Today*, *40*(2), 8–10. https://doi.org/10.1111/1467-8322.12872
Sigl, L. (2016). On the Tacit Governance of Research by Uncertainty: How Early Stage Researchers Contribute to the Governance of Life Science Research. *Science, Technology, & Human Values*, *41*(3), 347–374. https://doi.org/10.1177/0162243915599069
Sijtsma, K., Emons, W. H. M., Steneck, N. H., & Bouter, L. M. (2021). Steps toward preregistration of research on research integrity. *Research Integrity and Peer Review*, *6*(1), 5. https://doi.org/10.1186/s41073-021-00108-4
Simm, K., & Eigi-Watkin, J. (2024). Diverse sources of normativity in open science and their implications for ethical governance. *Royal Society Open Science*, *11*(7), 240480. https://doi.org/10.1098/rsos.240480
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/10.1177/0956797611417632
Simons, D. J. (2014). The Value of Direct Replication. *Perspectives on Psychological Science*, *9*(1), 76–80. https://doi.org/10.1177/1745691613514755
Sin, C. (2014). The policy object: A different perspective on policy enactment in higher education. *Higher Education*, *68*(3), 435–448. https://doi.org/10.1007/s10734-014-9721-5

Sismondo, S. (2021). Epistemic Corruption, the Pharmaceutical Industry, and the Body of Medical Science. *Frontiers in Research Metrics and Analytics*, *6*, 614013. https://doi.org/10.3389/frma.2021.614013

Sismondo, S. (2025). Cultures of Trial and Error: Self, Correction, and Science in Conflict-Laden Science. *Blog of Trial and Error*. https://doi.org/10.36850/c4b44bf2-c94a

Slaughter, S., & Leslie, L. L. (1997). *Academic capitalism: Politics, policies, and the entrepreneurial university*. Johns Hopkins University Press.

Smith, K. E. (2014). The politics of ideas: The complex interplay of health inequalities research and policy. *Science and Public Policy*, *41*(5), 561–574. https://doi.org/10.1093/scipol/sct085

Smith-Doerr, L., & Vardi, I. (2015). Mind the Gap: Formal Ethics Policies and Chemical Scientists' Everyday Practices in Academia and Industry. *Science, Technology, & Human Values*, *40*(2), 176–198. https://doi.org/10.1177/0162243914557950

Soderberg, C. K., Errington, T. M., Schiavone, S. R., Bottesini, J., Thorn, F. S., Vazire, S., Esterling, K. M., & Nosek, B. A. (2021). Initial evidence of research quality of registered reports compared with the standard publishing model. *Nature Human Behaviour*, *5*(8), 990–997. https://doi.org/10.1038/s41562-021-01142-4

Soehartono, A. M., Yu, L. G., & Khor, K. A. (2022). Essential signals in publication trends and collaboration patterns in global Research Integrity and Research Ethics (RIRE). *Scientometrics*, *127*(12), 7487–7497. https://doi.org/10.1007/s11192-022-04400-y

Sørensen, M. P., Lambert, M., Ravn, T., & Vostal, F. (2026). Time, Quality, and Integrity: Temporal Autonomy as a Missing Link in Research Assessment Reform. *Science and Engineering Ethics*, *32*(2), 16. https://doi.org/10.1007/s11948-026-00588-x

Sørensen, M. P., Ravn, T., Marušić, A., Elizondo, A. R., Kavouras, P., Tijdink, J. K., & Bendtsen, A.-K. (2021). Strengthening research integrity: Which topic areas should organisations focus on? *Humanities and Social Sciences Communications*, *8*(1), 198. https://doi.org/10.1057/s41599-021-00874-y

Spitzer, L., & Mueller, S. (2023). Registered report: Survey on attitudes and experiences regarding preregistration in psychological research. *PLOS ONE*, *18*(3), e0281086. https://doi.org/10.1371/journal.pone.0281086

Star, S. L. (2010). This is Not a Boundary Object: Reflections on the Origin of a Concept. *Science, Technology, & Human Values*, *35*(5), 601–617. JSTOR.

Star, S. L., & Griesemer, J. R. (1989). Institutional Ecology, `Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, *19*(3), 387–420. https://doi.org/10.1177/030631289019003001

Steneck, N. H. (1994). Research universities and scientific misconduct—History, policies, and the future. *The Journal of Higher Education*, *65*(3), 310–330.

Steneck, N. H. (2006). Fostering integrity in research: Definitions, current knowledge, and future directions. *Science and Engineering Ethics*, *12*(1), 53–74. https://doi.org/10.1007/PL00022268

Steneck, N. H. (2007). *ORI introduction to the responsible conduct of research* (Rev. ed.). U.S. Department of Health and Human Services, Office of Research Integrity. https://ori.hhs.gov/ori-introduction-responsible-conduct-research

Steneck, N. H. (2011). The dilemma of the honest researcher. *EMBO Reports*, *12*(8), 745–745. https://doi.org/10.1038/embor.2011.134

Steneck, N. H., Mayer, T., Anderson, M. S., & Kleinert, S. (2017). THE ORIGIN, OBJECTIVES, AND EVOLUTION OF THE WORLD CONFERENCES ON RESEARCH INTEGRITY. In L. C. Gundersen (Ed.), *Special Publications* (1st ed., pp. 1–14). Wiley. https://doi.org/10.1002/9781119067825.ch1

Stephan, P. (2012). Perverse incentives. *Nature*, *484*(7392), 29–31. https://doi.org/10.1038/484029a

Stephens, N., & Lewis, J. (2017). Doing laboratory ethnography: Reflections on method in scientific workplaces. *Qualitative Research*, *17*(2), 202–216. https://doi.org/10.1177/1468794116678040

Strathern, M. (Ed.). (2000). *Audit cultures: Anthropological studies in accountability, ethics, and the academy*. Routledge.

Stroebe, W., Postmes, T., & Spears, R. (2012). Scientific Misconduct and the Myth of Self-Correction in Science. *Perspectives on Psychological Science*, *7*(6), 670–688. https://doi.org/10.1177/1745691612460687

Stroebe, W., & Strack, F. (2014). The Alleged Crisis and the Illusion of Exact Replication. *Perspectives on Psychological Science*, *9*(1), 59–71. https://doi.org/10.1177/1745691613514450

Syed, M. (2025). Some data indicating that editors and reviewers do not check preregistrations during the review process. *Journal of Trial and Error*. https://doi.org/10.36850/e5ce-4cc5

Szollosi, A., & Donkin, C. (2021). Arrested Theory Development: The Misguided Distinction Between Exploratory and Confirmatory Research. *Perspectives on Psychological Science*, *16*(4), 717–724. https://doi.org/10.1177/1745691620966796

Szollosi, A., Kellen, D., Navarro, D. J., Shiffrin, R., Van Rooij, I., Van Zandt, T., & Donkin, C. (2020). Is Preregistration Worthwhile? *Trends in Cognitive Sciences*, *24*(2), 94–95. https://doi.org/10.1016/j.tics.2019.11.009

Tamot, R., Arsenieva, D., & Wright, D. E. (2013). The Emergence of the Responsible Conduct of Research (RCR) in PHS Policy and Practice. *Accountability in Research*, *20*(5–6), 349–368. https://doi.org/10.1080/08989621.2013.822258

Tang, B. L. (2025a). Publishing important work that lacks validity or reproducibility – pushing frontiers or corrupting science? *Accountability in Research*, *32*(7), 1159–1179. https://doi.org/10.1080/08989621.2024.2345714

Tang, G. (2025b). Mapping nine decades of research integrity studies (1935–2024): A scientometric analysis. *Accountability in Research*, 1–36. https://doi.org/10.1080/08989621.2025.2470860

Tavory, I. (2020). Interviews and Inference: Making Sense of Interview Data in Qualitative Research. *Qualitative Sociology*, *43*(4), 449–465. https://doi.org/10.1007/s11133-020-09464-x

The White House. (2025, May 23). Restoring Gold Standard Science. *The White House*. https://www.whitehouse.gov/presidential-actions/2025/05/restoring-gold-standard-science/

Thévenot, L. (2001). Pragmatic regimes governing the engagement with the world. *The Practice Turn in Contemporary Theory*.

Thompson, G. F., Supple-Green, J. F., Leva\vcić, R., & Mitchell, J. C. (1991). *Markets, Hierarchies and Networks: The Coordination of Social Life*. https://api.semanticscholar.org/CorpusID:141797853

Tijdink, J. K., Bouter, L. M., Veldkamp, C. L. S., Van De Ven, P. M., Wicherts, J. M., & Smulders, Y. M. (2016). Personality Traits Are Associated with Research Misbehavior in Dutch Scientists: A Cross-Sectional Study. *PLOS ONE*, *11*(9), e0163251. https://doi.org/10.1371/journal.pone.0163251

Tolich, M. (2004). Internal confidentiality: When confidentiality assurances fail relational informants. *Qualitative Sociology*, *27*(1), 101–106. https://doi.org/10.1023/B:QUAS.0000015546.20441.4a

Toth, A. A., Banks, G. C., Mellor, D., O'Boyle, E. H., Dickson, A., Davis, D. J., DeHaven, A., Bochantin, J., & Borns, J. (2021). Study Preregistration: An Evaluation of a Method for Transparent Reporting. *Journal of Business and Psychology*, *36*(4), 553–571. https://doi.org/10.1007/s10869-020-09695-3

Traweek, S. (1988). *Beamtimes and lifetimes: The world of high energy physicists*. Harvard
University Press.

Travis, G. D. L. (1981). Replicating Replication? Aspects of the Social Construction of Learning in Planarian Worms. *Social Studies of Science*, *11*(1), 11–32. https://doi.org/10.1177/030631278101100102

Troughton, L., & Obasi, A. (2022). An exploration of practices affecting research integrity in global health partnerships. *BMJ Global Health*, *7*(8), e009092. https://doi.org/10.1136/bmjgh-2022-009092

Tsai, A. C., Kohrt, B. A., Matthews, L. T., Betancourt, T. S., Lee, J. K., Papachristos, A. V., Weiser, S. D., & Dworkin, S. L. (2016). Promises and pitfalls of data sharing in qualitative research. *Social Science & Medicine*, *169*, 191–198. https://doi.org/10.1016/j.socscimed.2016.08.004

Tuffs, A. (2000). Fraud study concludes that self regulation has failed. *BMJ*, *321*(7253), 71A.

Tummers, L., Bekkers, V. J. J. M., & Steijn, B. (2012). *Policy Alienation of Public Professionals: A Comparative Case Study of Insurance Physicians and Secondary School Teachers*. (1 Online-Ressource (29 p)). SSRN. WorldCat. https://ssrn.com/abstract=2193279

Tyner, A. H., Abatayo, A. L., Daley, M., Field, S., Fox, N., Haber, N. A., Hahn, K. M., Struhl, M. K., Mawhinney, B., Miske, O., Silverstein, P., Soderberg, C. K., Stankov, T., Abbasi, A., Aberson, C. L., Aczel, B., Adamkovič, M., Albayrak, N., Allen, P. J., … Errington, T. M. (2026). Investigating the replicability of the social and behavioural sciences. *Nature*, *652*(8108), 143–150. https://doi.org/10.1038/s41586-025-10078-y

UK Committee on Research Integrity. (2024). *Enablers and Inhibitors of Research Integrity*. Zenodo. https://doi.org/10.5281/ZENODO.13891203

UK Research Integrity Office. (2025). *Code of Practice for Research: Promoting good practice and preventing misconduct (Version 3.5)*. UK Research Integrity Office. https://doi.org/10.37672/UKRIO.2023.04.codeofpractice

Ulpts, S. (2024). Responsible assessment of what research? Beware of epistemic diversity! *Meta-Psychology*, *8*. https://doi.org/10.15626/MP.2023.3797

Ulpts, S., Bartscherer, S. F., Field, S. M., & Penders, B. (2025). *The social replication of replication: Moving replication through epistemic communities*. https://doi.org/10.31235/osf.io/pqc4v_v2

Ulpts, S., & Schneider, J. W. (2024). *A conceptual review of uses and meanings of reproducibility and replication*. https://doi.org/10.31222/osf.io/entu4

Ulpts, S., & Schneider, J. W. (2025). *Knowledge Production Modes: The Relevance and Feasibility of Reproducibility*. MetaArXiv. https://doi.org/10.31222/osf.io/ujnd9_v2

Ulrich, R., & Miller, J. (2020). Questionable research practices may have little effect on replicability. *eLife*, *9*, e58237. https://doi.org/10.7554/eLife.58237

UNESCO. (2021). *UNESCO Recommendation on Open Science*. UNESCO. https://doi.org/10.54677/MNMH8546

US Public Health Service, (1989). Responsibilities of awardee and applicant institutions for dealing with and reporting possible misconduct in science; final rule. *Federal Register*, *54*(151), 32446–32451.

Vachon, B., Curran, J. A., Karunananthan, S., Brehaut, J., Graham, I. D., Moher, D., Sales, A. E., Straus, S. E., Fiander, M., Paprica, P. A., & Grimshaw, J. M. (2021). Replication Research Series-Paper 1: A concept analysis and meta-narrative review established a comprehensive theoretical definition of replication research to improve its use. *Journal of Clinical Epidemiology*, *129*, 176–187. https://doi.org/10.1016/j.jclinepi.2020.07.006

Valkenburg, G., Dix, G., Tijdink, J., & De Rijcke, S. (2021). Expanding Research Integrity: A Cultural-Practice Perspective. *Science and Engineering Ethics*, *27*(1), 10. https://doi.org/10.1007/s11948-021-00291-z

Van Den Akker, O. R., Bakker, M., Van Assen, M. A. L. M., Pennington, C. R., Verweij, L., Elsherif, M. M., Claesen, A., Gaillard, S. D. M., Yeung, S. K., Frankenberger, J.-L., Krautter, K., Cockcroft, J. P., Kreuer, K. S., Evans, T. R., Heppel, F. M., Schoch, S. F., Korbmacher, M., Yamada, Y., Albayrak-Aydemir, N., … Wicherts, J. M. (2024a). The potential of preregistration in psychology: Assessing preregistration producibility and preregistration-study consistency. *Psychological Methods*. https://doi.org/10.1037/met0000687

Van Den Akker, O. R., Van Assen, M. A. L. M., Bakker, M., Elsherif, M., Wong, T. K., & Wicherts, J. M. (2024b). Preregistration in practice: A comparison of preregistered and non-preregistered studies in psychology. *Behavior Research Methods*, *56*(6), 5424–5433. https://doi.org/10.3758/s13428-023-02277-0

Van Den Akker, O. R., Van Assen, M. A. L. M., Enting, M., De Jonge, M., Ong, H. H., Rüffer, F., Schoenmakers, M., Stoevenbelt, A. H., Wicherts, J. M., & Bakker, M. (2023). Selective Hypothesis Reporting in Psychology: Comparing Preregistrations and Corresponding Publications. *Advances in Methods and Practices in Psychological Science*, *6*(3), 25152459231187988. https://doi.org/10.1177/25152459231187988

Van Drimmelen, T., Slagboom, M. N., Reis, R., Bouter, L. M., & Van Der Steen, J. T. (2024). Decisions, Decisions, Decisions: An Ethnographic Study of Researcher Discretion in Practice. *Science and Engineering Ethics*, *30*(6), 59. https://doi.org/10.1007/s11948-024-00481-5

Van Drimmelen, T., Slagboom, M. N., Reis, R., Bouter, L. M., & Van Der Steen, J. T. (2025). Navigating Ambiguity and Uncertainty in End-of-Life Care Research Practice. *Journal of Empirical Research on Human Research Ethics*, *20*(5), 323–334. https://doi.org/10.1177/15562646251377121

van Elk, M., Rowatt, W., & Streib, H. (2018). Good Dog, Bad Dog: Introducing Open Science Badges. *The International Journal for the Psychology of Religion*, *28*(1), 1–2. https://doi.org/10.1080/10508619.2018.1402589

Van Noorden, R. (2023). More than 10,000 research papers were retracted in 2023—A new record. *Nature*, *624*(7992), 479–481. https://doi.org/10.1038/d41586-023-03974-8

van Rooij, I. (2019, January 18). Psychological science needs theory development before preregistration. *Psychonomic Society Featured Content*. https://featuredcontent.psychonomic.org/psychological-science-needs-theory-development-before-preregistration/

Vasconcelos, S., & Marušić, A. (2025). Gen AI and research integrity: Where to now?: The integration of Generative AI in the research process challenges well-established definitions of research integrity. *EMBO Reports*, *26*(8), 1923–1928. https://doi.org/10.1038/s44319-025-00424-6

Vazire, S. (2018). Implications of the Credibility Revolution for Productivity, Creativity, and Progress. *Perspectives on Psychological Science*, *13*(4), 411–417. https://doi.org/10.1177/1745691617751884

Vazire, S. (2024). The Next Chapter for *Psychological Science*. *Psychological Science*, *35*(7), 703–707. https://doi.org/10.1177/09567976231221558

Vendé, B., Barberousse, A., & Ruphy, S. (2025). From 2015 to 2023, eight years of empirical research on research integrity: A scoping review. *Research Integrity and Peer Review*, *10*(1), 5. https://doi.org/10.1186/s41073-025-00163-1

Vertesi, J. (2020). *The Social Life of Spacecraft: Organizations, Decisions, and Science on NASA's Robotic Teams*. University of Chicago Press. https://doi.org/10.7208/chicago/9780226691114.001.0001

Videnoja, K., Tauginienė, L., & Löfström, E. (2024). Family without kinship – the pluralism of European regulatory research integrity systems and its implications. *Accountability in Research*, 1–22. https://doi.org/10.1080/08989621.2024.2345710
Vincent, M., & Hooper, M. (2025). Research integrity, thick and thin. *Synthese*, *206*(2), 79. https://doi.org/10.1007/s11229-025-05157-6
Von Bargen, J. (2013). Wissenschaftliche Redlichkeit und zentrales hochschulinternes Verfahrensrecht. *JuristenZeitung*, *68*(14), 714–723. https://doi.org/10.1628/002268813X13710404485887
Wagenmakers, E.-J., Wetzels, R., Borsboom, D., Van Der Maas, H. L. J., & Kievit, R. A. (2012). An Agenda for Purely Confirmatory Research. *Perspectives on Psychological Science*, *7*(6), 632–638. https://doi.org/10.1177/1745691612463078
Wäscher, S., Biller-Andorno, N., & Deplazes-Zemp, A. (2020). "I Don't Want to Do Anything Bad." Perspectives on Scientific Responsibility: Results from a Qualitative Interview Study with Senior Scientists. *NanoEthics*, *14*(2), 135–153. https://doi.org/10.1007/s11569-020-00365-5
WCRI. (2025, September 1). *Amsterdam Agenda*. WCRIF - The World Conferences on Research Integrity Foundation. https://www.wcrif.org/guidance/amsterdam-agenda
Weingart, P. (2017). *Is There a Hype Problem in Science? If So, How Is It Addressed?* (K. H. Jamieson, D. M. Kahan, & D. A. Scheufele, Eds.; Vol. 1). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780190497620.013.12
Whitley, R. (1984). *The intellectual and social organization of the sciences*. Clarendon Press ; Oxford University Press.
Whitley, R. D. (1970). Black Boxism and the Sociology of Science: A Discussion of the Major Developments in the Field. *The Sociological Review*, *18*(1_suppl), 61–92. https://doi.org/10.1111/j.1467-954X.1970.tb03176.x
Whitley, R. D. (1983). From the sociology of scientific communities to the study of scientists' negotiations and beyond. *Social Science Information*, *22*(4–5), 681–720. https://doi.org/10.1177/053901883022004004
Wiggins, B. J., & Christopherson, C. D. (2019). The replication crisis in psychology: An overview for theoretical and philosophical psychology. *Journal of Theoretical and Philosophical Psychology*, *39*(4), 202–217. https://doi.org/10.1037/teo0000137
Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18
Willig, C. (2013). *Introducing qualitative research in psychology* (3rd ed). McGraw-Hill Open University Press.

Willig, C., & Stainton-Rogers, W. (2017). *The SAGE handbook of qualitative research in psychology* (Second edition). SAGE.
Winchester, C. (2018). Give every paper a read for reproducibility. *Nature*, *557*(7705), 281–281. https://doi.org/10.1038/d41586-018-05140-x
Wolfe, A. J. (2018). *Freedom's Laboratory: The Cold War Struggle for the Soul of Science*. Johns Hopkins University Press.
Woo, S., & Walsh, J. P. (2024). On the shoulders of fallen giants: What do references to retracted research tell us about citation behaviors? *Quantitative Science Studies*, *5*(1), 1–30. https://doi.org/10.1162/qss_a_00303
Wouters, P. (1997). Citation cycles and peer review cycles. *Scientometrics*, *38*(1), 39–55. https://doi.org/10.1007/BF02461122
Yadav, D. (2022). Criteria for Good Qualitative Research: A Comprehensive Review. *The Asia-Pacific Education Researcher*, *31*(6), 679–689. https://doi.org/10.1007/s40299-021-00619-0
Yamada, Y. (2018). How to Crack Pre-registration: Toward Transparent and Open Science. *Frontiers in Psychology*, *9*, 1831. https://doi.org/10.3389/fpsyg.2018.01831
Yarkoni, T. (2018, October 2). *No, it's not The Incentives—It's you* [Tumblr]. [Citation Needed] and after the Citation, Some Ice Cream Would Be Nice. https://embed.tumblr.com/widgets/share/button?canonicalUrl=https%3A%2F%2Ftalyarkoni.org%2Fblog%2F2018%2F10%2F02%2Fno-its-not-the-incentives-its-you%2F&postcontent%5Btitle%5D=No%2C%20it%27s%20not%20The%20Incentives%E2%80%94it%27s%20you&postcontent%5Bcontent%5D=https%3A%2F%2Ftalyarkoni.org%2Fblog%2F2018%2F10%2F02%2Fno-its-not-the-incentives-its-you%2F
Yeo-Teh, N. S. L., & Tang, B. L. (2024). On "intent" in research misconduct. *Accountability in Research*, 1–19. https://doi.org/10.1080/08989621.2024.2374577
Zhaksylyk, A., Zimba, O., Yessirkepov, M., & Kocyigit, B. F. (2023). Research Integrity: Where We Are and Where We Are Heading. *Journal of Korean Medical Science*, *38*(47), e405. https://doi.org/10.3346/jkms.2023.38.e405
Zuckerman, H. (1984). Norms and Deviant Behavior in Science. *Science, Technology, & Human Values*, *9*(1), 7–13. https://doi.org/10.1177/016224398400900102
Zuckerman, H. (2020). Is "the time ripe" for quantitative research on misconduct in science? *Quantitative Science Studies*, *1*(3), 945–958. https://doi.org/10.1162/qss_a_00065
Zuckerman, H., & Sagarin, E. (1977). Deviance and social change. In *Deviance and social change* (pp. 87-138). Sage Beverly Hills.
Zwaan, R. A., Etz, A., Lucas, R. E., & Donnellan, M. B. (2018). Making replication mainstream. *Behavioral and Brain Sciences*, *41*, e120. https://doi.org/10.1017/S0140525X17001972

Appendices

Appendix A

Interview guide

Einführung

Einwilligung

Kannst du mir ein bisschen erzählen woran du in letzter Zeit so gearbeitet hast?

Forschungsbeobachtungen

Was würdest du sagen woher weißt du was gute Forschung ist, woran orientierst du dich?

Gibt es aufgaben oder Situationen als Wissenschaftler (Peer Review, Proposal schreiben, Artikel schreiben etc.) wo du denkst du könntest mehr Führung oder Orientierung gebrauchen?

Du hast öfter mal erwähnt, dass du viel zu tun hast und daher manchmal wenig Zeit. Daher wollte Ich fragen, welche Rolle spielt der Faktor Zeit, gerade wenn es darum geht gut zu forschen?

Du bist ja auch Open Science Ambassador, warum ist Open Science wichtig für eure Forschung.

Du hattest auch einmal erwähnt, dass Open Science gute Wissenschaft ist. Kannst du das ein bisschen ausführen?

Ein Teil deiner Arbeit ist ja auch Metascience, kannst du mir sagen, warum du an Metascience arbeitest?

Etwas worauf ihr auch sehr Wert zu legen scheint ist Replikation und Reproduzierbarkeit. Kannst du mir sagen, was das für dich bedeutet und warum das für eure Forschung wichtig ist?

Etwas anderes, was Ich beobachtet habe war bei einem der Colloquien hat einer von euch darauf hingewiesen, dass man auch angeben muss dass das Ergebnis ein

‚Statistical fluke' sein kann. Später habe Ich dann gehört, dass man sowas in anderen Gruppen oder Abteilung nie erwähnen würde. Wieso ist das so wichtig solche Umstände preiszugeben?

Was siehst du als Verantwortung als Supervisor wenn es um gute oder verantwortungsvolle Forschung geht?

Was denkst du sind Probleme in deiner Disziplin?
    Wieso sind das Probleme?
    Was macht diese Probleme wichtig?

Forschungsintegrität

Hattest du einen Kurs oder Lehre in guter wissenschaftlicher Praxis oder so etwas?

Redet ihr in der Forschungsgruppe über solche Sachen miteinander?

Gibt es in deinem Forschungsinstitut oder deiner direkten Forschungsumgebung etwas das du ändern würdest oder wo du sagst, hier ist etwas das mich an der guten Ausführung meiner Forschung behindert?

Wenn du den Begriff Forschungsintegrität hörst oder liest was stellst du dir darunter vor?

Interview guide

Consent

You read the study information and consent to do the interview?

To start, can you tell me a bit about what you have been working on recently?

Field observations

Can you tell me a bit what you see as your role and responsibility as a lab leader or PI when it comes to responsible/ good research?

Similarly, can you tell me a bit what you see as your role and responsibility as a supervisor when it comes to responsible/ good research?

During meetings you asked others whether they preregistered their study. Why do you ask that?

During one of the meetings, you also mentioned making data and code openly available for others to see and use. Why do you do that? Why is that important for you or your research?

Sometimes the issue of reproducibility or replication came up. Can you tell me what those concepts mean to you (what is that) and why it is important for your research?

Then I also noticed that the factor time comes up when you mentioned being behind on your work or still having to do 10 reviews. How would you describe the relation between workload and available time with good or responsible practice?

> Can you describe situation when workload or lack of time gets in the way of good practice?

I also often came across the question: What is novel about it or what is the novel contribution? Why is that so important?

During one of your meetings, you were talking about reviewer recommendations for a journal submission. You mentioned that it would be okay to recommend someone whose dataset you use as long as you haven't published with them and they don't work in the department (anymore). What goes into the decision process of deciding who to recommend?

> How do you know who is appropriate to recommend? Or the other way around how do you know what is inappropriate?

During a reading group when you shortly mentioned the reform movement you said that of course you want good statistics, but at the same time you also want to have space in the paper for the actual topic of the paper. Can you elaborate a bit on that issue?

Interesting were also short discussions about authorship and who to put on a paper as an author. How do you decide who to list as an author?

How do you know how to make that decision and what to consider?

Another thing I noticed is that the topic of money or funding comes up quite a bit. How would you say financing or funding availability relates to being able to do good research?

In previous labs some also mentioned the aspect of the research climate (how you get along how much pressure you feel). How does that influence your research?

You also mentioned the grant or funding application procedure and that you don't have a talent for it. Can you elaborate on that? What is it that makes it special or problematic for you?

You also had job talks; how do you assess them? Or in other words, what is it that you look out for in applicants and why?

Research Integrity

Did you have a course or something in good scientific practice or responsible research conduct?

How do you in the group talk about such things?

Are there things in the institute or the direct environment that you would change or where you think this is something that makes it harder for me to conduct good research?

When you read or hear the term research integrity, what comes to mind?

Appendix B

From the researcher to the integrity of knowledge production – all codes

Name	Description
Preregistration	Captures how researchers talk about the practice of preregistration as a component of good research.
Credibility Management Strategy	This captures how the researchers talk about using preregistration as a reporting tool to persuade their research community that their research and they are credible.
Constrain Researcher Degrees of Freedom	Captures how researchers talk about limiting the effect of degrees of freedom on their research, especially due to the sensitivity of results to said degrees of freedom.
Thinking through your Study	Captures how preregistration can help researchers in properly planning and designing their study, for example, to make it a stronger test of their hypothesis.
Time Stamped Firewall between confirmation and exploration	Captures how researchers report that preregistration helps them in communicating to their community the order of their conduct and which parts of the research are confirmatory and which are exploratory.
Hardship of Preregistration	This captures the practicalities and problems of practicing preregistration, that the researchers reflect on.
Investment & Workload	This captures the necessary investment and workload that the researchers associate with practicing preregistration.
Maybe not a Prison, but a Punishment	Captures the researchers' perception that they do not necessarily feel restricted in their conduct, when the study is preregistered, but they do feel like an extra burden is added, for instance, when they deviate from the preregistration.
Overwhelming student	Captures their opinion that for preregistration a certain amount of expertise is necessary which students do not have. This also includes potential for exploiting students.
Uncertainty of Piloting	This captures the perceived lack of guidance and uncertainty related to preregistration in the context of piloting.
Replication & Reproducibility	Captures the researchers' discourse about the practices of replication and reproducibility and how they relate to good research.
Functions	Captures what the scholars say is the utility they ascribe to the practices of replication and reproducibility for their research.
Stable foundation	Captures how the researchers talk about how replication and reproducibility help them in building their own research on a stable foundation.
Meanings	Captures how the researchers use and define the terms replication and reproducibility for themselves.
Confusion	Captures how the talk from the researchers demonstrates confusion about the meanings of reproducibility and replication.

Name	Description
Stories from their own Practice	Captures how and which forms of replication and reproducibility they employ in their own and routines.
Failed Replication	Captures how they talk about navigating and dealing with failed replications.
Novelty Incentives	Captures how they talk about how demands and incentives for novelty, set by the wider science system and their institute, affect practicing replication and reproducibility.
Research Ethics	Captures the researchers' discourse about the ethicality of their conduct and how it relates to 'good research'.
Dealing with Participants	Captures how the researchers talk about the relationship between their ethicality of conduct in interaction with participants and research integrity ('good research').
The ethical Colleague	Captures how the researchers talk about the relationship between the ethicality of conduct in interaction with colleagues in the lab and research integrity ('good research').
Research Integrity	Captures the researchers' talk about the meaning of research integrity.
Moral Scholar & Personal Integrity	Captures how the morality and personality of the researchers, as human beings, relate to research integrity as well as their notions of the researchers' responsibility for research integrity.
Honesty	Captures how they relate being an honest person and researchers to research integrity.
Humility & Critical Reflexivity	Captures the relationship of being humble, reflexive and (self-)critical to research integrity, as 'good research'
Taking Responsibility & owning the Research	Captures how the act of taking responsibility and taking ownership for one's research, especially in the face of scrutiny, relates to research integrity and 'good research'.
Term-Concept not known	Captures the researchers reports about a lack of familiarity with the concept or term 'research integrity' and how they find it foreign, strange, or weird.
The Integrity of the Research	Captures the characteristics of the research that the scholars relate to research integrity ('good research').
Methodological Rigor	Captures the researchers' discourse about the centrality and role of proper and correct application of methods for research integrity ('good research').
Hypothesis Testing	Captures the reported centrality of hypothesis driven research for research integrity ('good research').
Prevention of Bias	Captures their talk about making sure that bias or uncontrolled researcher intervention did not affect their research and findings, as well as how that relates to research integrity and 'good research'.
Research Ethics (do no harm)	Captures their discourse about navigating procedural ethics protocols as part of their research and how that relates to research integrity ('good research').

Name	Description
Misconduct	Captures the researchers' discourse that relates research integrity directly to misconduct.
Societal Relevance & Practical Utility	Captures how they talk about that the societal relevance and practical utility of their research is or should be a part of research integrity.
Research Integrity & Open Science	Captures how the researchers relate Open Science to Research Integrity ('good research').
The Researcher in a Group of Researchers	Captures how the researchers talk about how the mutual dependence between researchers and the role of trust is a component of research integrity, or a condition for practicing 'good research'.
The Role of the Research System	Captures how the researchers talk about the demands and incentives of the wider research system, or their institutes and how that relates to research integrity and their ability to conduct 'good research'.